GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:46:58 ; Search time 209 Seconds
(without alignments)
872.784 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 710428

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113320_7333/app_query.fasta_1.263
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_208_@runat_07052003_113320_7333 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		56	12.7	42	24	AAD27031	Human Secreted epi
2	c	49	11.1	50	8	AAN70336	3'-5' sequence of
3		47	10.6	51	22	AAI78209	Human silent SNP c
4	c	47	10.6	60	24	ABN39530	Human spliced tran
5	c	46	10.4	60	24	ABN45376	Human spliced tran
6	c	46	10.4	65	24	ABN52611	Mouse spliced tran
7		46	10.4	66	24	ABN73405	Bovine embryonic g
8		46	10.4	68	20	AAX35998	Polynucleotide ass
9		46	10.4	71	21	AAA49245	HCV core protein p
10		46	10.4	71	21	AAA46147	Plasmid pgHCV/delt
11	c	46	10.4	72	14	AAQ51446	Human FACC intron
12		46	10.4	79	14	AAQ36160	Mutagenic primer C
13	c	45	10.2	48	22	AAI76614	Human silent SNP c
14		44.5	10.1	65	24	ABN31870	Rat spliced transc
15	c	44	10.0	50	15	AAQ54486	Excitatory amino a
16	c	44	10.0	50	15	AAQ54487	Excitatory amino a
17		44	10.0	60	21	AAA50222	Hybrid leptin/tiss
18		44	10.0	63	21	AAA50221	Hybrid leptin/tiss
19		44	10.0	75	22	AAD09429	V1Jns-tPA vaccine
20	c	43.5	9.8	65	24	ABN31369	Rat spliced transc
21		43.5	9.8	76	21	AAA99033	Human TGC839 PCR p
22		43.5	9.8	80	24	ABL59337	PCR primer for hum
23		43	9.7	50	22	AAL30844	Human SNP oligonuc
24		43	9.7	51	22	AAI75717	Human silent SNP c
25		43	9.7	60	24	ABN38069	Human spliced tran
26	c	43	9.7	60	24	ABN40440	Human spliced tran
27	c	43	9.7	65	24	ABN31099	Rat spliced transc
28		43	9.7	66	21	AAA49049	DNA encoding tissu
29		43	9.7	66	21	AAA50218	Modified human tis
30		43	9.7	66	22	AAC66513	Tissue plasminogen
31		43	9.7	69	21	AAA50217	Native human tissu
32		43	9.7	69	21	AAA50219	Modified human tis
33		43	9.7	69	21	AAA50220	Modified human tis
34		43	9.7	69	21	AAA09463	tPA leader sequenc
35		43	9.7	69	24	ABL39945	Signal peptide Tpa
36	c	43	9.7	70	23	ABL58043	Cyclooxygenase-2,
37		43	9.7	75	24	ABL39944	Signal peptide Tpa
38		43	9.7	76	21	AAA50338	Synthetic oligonuc
39		43	9.7	78	17	AAT15476	Human tissue-speci
40		43	9.7	78	18	AAT85512	Human tissue-speci
41		43	9.7	78	19	AAV21738	Sense oligomer 1 u
42	c	43	9.7	78	19	AAV21739	Antisense oligomer
43		43	9.7	78	22	AAS09203	Synthetic sense ol
44	c	43	9.7	78	22	AAS09204	Synthetic antisens
45	c	43	9.7	78	22	AAD09430	V1Jns-tPA vaccine

ALIGNMENTS

RESULT 1
AAD27031
ID AAD27031 standard; DNA; 42 BP.
XX
AC AAD27031;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 DNA amplifying PCR primer #5.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human; PCR primer; ss.
XX

```
OS   Homo sapiens.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Example 4; Page 89; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   DNA amplifying PCR primer.
XX
SQ   Sequence 42 BP; 15 A; 10 C; 9 G; 8 T; 0 other;

Alignment Scores:
Pred. No.:              224          Length:        42
Score:                  56.00        Matches:       10
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            12.67%       Indels:        0
DB:                     24           Gaps:          0

US-09-599-087B-5 (1-81) x AAD27031 (1-42)

Qy      25 LysArgArgProAlaLysAlaTrpSerGly 34
           ||||||||||||||||||||||||||||||
Db      13 AAACGTCGTCCAGCTAAAGCCTGGTCAGGC 42

RESULT 2
AAN70336/c
ID   AAN70336 standard; DNA; 50 BP.
XX
AC   AAN70336;
XX
DT   03-OCT-2002  (updated)
DT   02-APR-1991  (first entry)
XX
DE   3'-5' sequence of probe A11 used to detect hepatitis B virus (HBV)
DE   analyte.
XX
KW   Hepatitis B virus assay; diagnosis; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..20
FT                   /*tag=  a
FT                   /note="Fluorescein label conjugate binding site"
FT   misc_feature    21..50
FT                   /*tag=  b
FT                   /note="probe segment to HBV"
XX
PN   EP225807-A.
XX
PD   16-JUN-1987.
XX
PF   10-DEC-1986;   86EP-0309622.
XX
PR   11-DEC-1985;   85US-0807624.
PR   23-DEC-1986;   86US-0945876.
XX
PA   (CHIR-) CHIRON CORP.
XX
PI   Ureda MS,  Warner B,  Horn J;
XX
DR   WPI; 1987-165004/24.
XX
PT   Detecting specific nucleic acid sequences - by incubating sample
PT   with labelling reagent set and capturing reagent set.
XX
PS   Example; page 17; 31pp; English.
XX
CC   In the example a BglII fragment of hepatitis B virus (AAN70325) was
CC   detected.  12 sequences, complimentary to different sequences
CC   present in HBV were constructed (see AAN70326-n70337).  Six of the HBV
CC   complementary sequences are joined to a common sequence (A) for
CC   complexing with the label conjugate (A').  The other six HBV
CC   complementary sequences are joined to a common sequence (B) for
CC   complexing with a biotinylated sequence (B') or a third DNA sequence
CC   (B'C) for binding to a support (see AAN70338, AAN70339, AAN70340).
CC   (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ   Sequence 50 BP; 13 A; 8 C; 14 G; 15 T; 0 other;

Alignment Scores:
Pred. No.:              1.56e+03     Length:        50
Score:                  49.00        Matches:       8
Percent Similarity:     78.57%       Conservative:  3
Best Local Similarity:  57.14%       Mismatches:    3
Query Match:            11.09%       Indels:        0
DB:                     8            Gaps:          0

US-09-599-087B-5 (1-81) x AAN70336 (1-50)

Qy      40 CysCysHisArgValProSerProAsnSerThrAsnLeuLys 53
           |||||||||:::   ||||||   :::   |||||||||:::
Db      45 TGTTGTCATCAAAGGCCTTCACAACACTTGACCAACTTGCAA 4

RESULT 3
AAI78209
ID   AAI78209 standard; DNA; 51 BP.
XX
AC   AAI78209;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:5150.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
```

XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 2086; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 14 A; 13 C; 13 G; 11 T; 0 other;

Alignment Scores:

Pred. No.:	2.63e+03	Length:	51
Score:	47.00	Matches:	8
Percent Similarity:	72.73%	Conservative:	0
Best Local Similarity:	72.73%	Mismatches:	3
Query Match:	10.63%	Indels:	0
DB:	22	Gaps:	0

US-09-599-087B-5 (1-81) x AAI78209 (1-51)

```
Qy     39 LeuCysCysHisArgValProSerProAsnSer 49
          |||||||||   |||   ||||||   ||||||
Db     19 TTGTGCTGTGAACGAGCACCGTCAGAGAATTCC 51
```

RESULT 4
ABN39530/c
ID ABN39530 standard; DNA; 60 BP.
XX
AC ABN39530;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:12278.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 12278; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 13 A; 23 C; 9 G; 15 T; 0 other;

Alignment Scores:

Pred. No.:	3.23e+03	Length:	60
Score:	47.00	Matches:	9
Percent Similarity:	71.43%	Conservative:	1
Best Local Similarity:	64.29%	Mismatches:	4
Query Match:	10.63%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN39530 (1-60)

```
Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
          |||:::||||||   |||||||||         |||||||||
Db     56 TCGTCTGAAGGCCTAAGGAGACCTTGGCTTGGATGGAGTGGG 15
```

RESULT 5
ABN45376/c
ID ABN45376 standard; DNA; 60 BP.
XX
AC ABN45376;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:18124.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX

PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 18124; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 16 C; 21 G; 9 T; 0 other;

Alignment Scores:

Pred. No.:	4.13e+03	Length:	60
Score:	46.00	Matches:	7
Percent Similarity:	63.64%	Conservative:	0
Best Local Similarity:	63.64%	Mismatches:	4
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN45376 (1-60)

```
Qy     62 ProCysLysLeuGluProGluProArgLeuTrp 72
          ||||||   |||   |||   ||||||   |||
Db     55 CCGTGCAGTCTCGTTCCGAGCCCGAGATCCTGG 23
```

RESULT 6
ABN52611/c
ID ABN52611 standard; DNA; 65 BP.
XX
AC ABN52611;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:25359.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 25359; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 20 A; 16 C; 13 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	4.58e+03	Length:	65
Score:	46.00	Matches:	8
Percent Similarity:	78.57%	Conservative:	3
Best Local Similarity:	57.14%	Mismatches:	3
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN52611 (1-65)

```
Qy     38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsn 51
          |||:::|||   ||||||:::||||||   :::|||   |||
Db     61 AGAGTTTGCACTCATAGGATCCCAAGTGGTTCTTCATAAAAT 20
```

RESULT 7
ABN73405
ID ABN73405 standard; cDNA; 66 BP.
XX
AC ABN73405;
XX
DT 03-JUL-2002 (first entry)

XX
DE Bovine embryonic germ (EG) cell cDNA EST 990729a CONTIG 2 #4.
XX
KW Bovine; Bos taurus; EST; expressed sequence tag; totipotence;
KW development; gene; ss.
XX
OS Bos taurus.
XX
PN WO200194550-A2.
XX
PD 13-DEC-2001.
XX
PF 07-JUN-2001; 2001WO-US18576.
XX
PR 07-JUN-2000; 2000US-209874P.
PR 06-JUN-2001; 2001US-0876143.
XX
PA (INFI-) INFIGEN INC.
XX
PI Eilertsen KJ, Pfister-Genskow M, Childs L;
XX
DR WPI; 2002-351289/38.
XX
PT An expressed sequence tag (EST), the expression of which, or its
PT complementary sequence, in a cell identifies the cell as a
PT developmentally competent or incompetent cell -
XX
PS Example 16; Page 165; 584pp; English.
XX
CC The present invention describes an expressed sequence tag (EST), where
CC the EST is an isolated, enriched, or purified nucleic acid sequence
CC representing all or part of a gene, the expression of which, or its
CC complementary sequence, in a cell identifies the cell as a
CC developmentally competent or incompetent cell. Molecules which induce
CC developmental competence in a cell line are useful for inducing
CC totipotence in one or more cells. Molecules which induce developmental
CC incompetence in a cell line are useful for preventing a full term
CC pregnancy in an animal and inhibiting totipotence. The molecules are
CC also useful for treating a disease in an animal by inducing development
CC of one or more cells of the animal into a specific cell type. The
CC present sequence represents a bovine EST which is given in the
CC exemplification of the present invention.
XX
SQ Sequence 66 BP; 4 A; 12 C; 6 G; 44 T; 0 other;

Alignment Scores:

Pred. No.:	4.67e+03	Length:	66
Score:	46.00	Matches:	9
Percent Similarity:	66.67%	Conservative:	3
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN73405 (1-66)

```
Qy       3 LeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||      :::||||||   ||||||:::|||   |||   |||   :::|||
Db      10 CTTTTTTATATCTCTTCCTTCCTTTGTGTTTTGTGTCTCTTTTTTTTTCTCTTT 63
```

RESULT 8
AAX35998
ID AAX35998 standard; DNA; 68 BP.
XX
AC AAX35998;
XX
DT 15-JUL-1999 (first entry)
XX
DE Polynucleotide associated with the activation of neutrophils.
XX
KW Neutrophil; differentially expressed; bacteria; activation;
KW nucleic acid probe; amplification primer; hybridisation assay;
KW pathogen; ss.
XX
OS Homo sapiens.
XX
PN WO9920642-A2.
XX
PD 29-APR-1999.
XX
PF 21-OCT-1998; 98WO-US22171.
XX
PR 21-OCT-1997; 97US-0062564.
XX
PA (UYYA) UNIV YALE.
XX
PI Goguen J, Newburger P, Prashar Y, Weissman S, Yamaga S;
PI Yerramilli S;
XX
DR WPI; 1999-288271/24.
XX
PT Unique nucleic acid sequence
XX
PS Claim 1; Page 51; 74pp; English.
XX
CC AAX35979-X36064 represents neutrophil nucleic acid sequences which
CC correspond to mRNA, and therefore to genes. These sequences are
CC differentially expressed upon exposure of the neutrophils to virulent
CC or avirulent bacteria. The nucleic acid sequences can be used as
CC nucleic acid probes to screen cDNA and genomic libraries to select
CC similar sequences; to generate amplification primers; and to identify
CC adjacent and flanking sequences in the genome. The sequences can be
CC used, e.g. in hybridisation assays, to detect activated neutrophils or
CC neutrophils previously exposed to virulent or avirulent bacteria. The
CC sequences can also be used as probes to monitor the expression of at
CC lest one differentially expressed neutrophil gene in a method to
CC identify a therapeutic or prophylactic agent that modulates the response
CC of a neutrophil to a pathogen.
XX
SQ Sequence 68 BP; 17 A; 7 C; 25 G; 17 T; 2 other;

Alignment Scores:

Pred. No.:	4.85e+03	Length:	68
Score:	46.00	Matches:	7
Percent Similarity:	63.64%	Conservative:	0
Best Local Similarity:	63.64%	Mismatches:	4
Query Match:	10.41%	Indels:	0
DB:	20	Gaps:	0

US-09-599-087B-5 (1-81) x AAX35998 (1-68)

```
Qy      31 AlaTrpSerGlyArgArgThrArgLeuCysCys 41
             ||||||   |||||||||         ||||||
Db       2 GCCTGGGANGGGAGGAGAGGAGAGNCGTGCTGT 34
```

RESULT 9
AAA49245
ID AAA49245 standard; DNA; 71 BP.
XX
AC AAA49245;
XX
DT 19-DEC-2000 (first entry)
XX
DE HCV core protein plasmid pgHCVdelta145-154 mutagenic oligonucleotide #12.
XX
KW Hepatitis C; targeted protein production; core protein;
KW mutagenic primer; ss.
XX
OS Synthetic.
XX
PN WO200031127-A2.
XX
PD 02-JUN-2000.
XX
PF 23-NOV-1999; 99WO-GB03898.
XX
PR 26-NOV-1998; 98GB-0025953.

XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Hope G, Mclauchlan J;
XX
DR WPI; 2000-400040/34.
XX
PT Protein comprising a lipid globule targeting sequence consisting of a
PT hepatitis C virus core protein, useful for targeting a protein of
PT interest to lipid globules which are subsequently secreted into animal
PT milk -
XX
PS Example; Page 32; 60pp; English.
XX
CC The present sequence is a mutagenic primer used in the construction of a
CC plasmid containing a mutated version of the hepatitis C virus core
CC protein gene. The core protein has been shown to be localised to lipid
CC droplets within cells. This means that fragments of the protein, in
CC particular those comprising amino acids 125-144 or 161-166, can be used
CC to target proteins to lipid droplets in milk. This provides a simple and
CC efficient means of obtaining proteins of interest, using transgenic
CC animals. Proteins of interest include those implicated in disease and
CC antigenic polypeptides for use as vaccines.
XX
SQ Sequence 71 BP; 13 A; 20 C; 22 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	5.13e+03	Length:	71
Score:	46.00	Matches:	9
Percent Similarity:	68.75%	Conservative:	2
Best Local Similarity:	56.25%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	21	Gaps:	0

US-09-599-087B-5 (1-81) x AAA49245 (1-71)

```
Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66
```

RESULT 10
AAA46147
ID AAA46147 standard; cDNA; 71 BP.
XX
AC AAA46147;
XX
DT 27-OCT-2000 (first entry)
XX
DE Plasmid pgHCV/delta145-154 insertion oligonucleotide.
XX
KW Hepatitis C virus; core protein; HCV; lipid globule; chronic hepatitis;
KW liver disease; ADRP displacement; insertion oligonucleotide; ss.
XX
OS Synthetic.
XX
PN WO200031532-A1.
XX
PD 02-JUN-2000.
XX
PF 23-NOV-1999; 99WO-GB03906.
XX
PR 26-NOV-1998; 98GB-0025951.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Hope G, Mclauchlan J;
XX
DR WPI; 2000-400181/34.
XX
PT Identifying a substance for treating or preventing a viral infection
PT such as hepatitis C virus, comprises determining if the substance
PT disrupts a lipid globule target sequence to lipid globule interaction
PT -
XX
PS Examples; Page 37; 74pp; English.
XX
CC The present sequence is an oligonucleotide which was inserted into
CC a plasmid which contained part of hepatitis c virus core protein gene.
CC The virus causes chronic hepatitis and liver disease in humans. It is
CC thought that it does this by associating with intracellular lipid
CC droplets and downregulating the expression of adipocyte-specific
CC differentiation-related protein (ADRP), which is probably required for
CC the maintenance of lipid droplets. The core protein can be used to
CC identify substances capable of disrupting this interaction, which are
CC thus capable of treating or preventing infection by the hepatitis c virus
CC and its accompanying effects. This oligonucleotide was used in the
CC construction of mutants of the core protein, which were used to determine
CC their efficiency of infection compared to the wild-type.
XX
SQ Sequence 71 BP; 13 A; 20 C; 22 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	5.13e+03	Length:	71
Score:	46.00	Matches:	9
Percent Similarity:	68.75%	Conservative:	2
Best Local Similarity:	56.25%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	21	Gaps:	0

US-09-599-087B-5 (1-81) x AAA46147 (1-71)

```
Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66
```

RESULT 11
AAQ51446/c
ID AAQ51446 standard; DNA; 72 BP.
XX
AC AAQ51446;
XX
DT 20-MAY-1994 (first entry)
XX
DE Human FACC intron 9 3' region.
XX
KW Fanconi Anemia Group C; FACC; complementing cDNA; varient; diagnosis;
KW open reading frame; Fanconi anemia; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9322435-A.
XX
PD 11-NOV-1993.
XX
PF 27-APR-1993; 93WO-CA00178.
XX
PR 29-APR-1992; 92US-0876285.
PR 21-JUL-1992; 92US-0918313.
PR 15-JAN-1993; 93US-0003963.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
XX
PI Buchwald M, Mathew CG, Strathdee CA, Wevrick R;
XX
DR WPI; 1993-368794/46.
XX
PT Human cDNA which complements Fanconi Anaemia gp. C - used to
PT develop prods. for use in diagnosis, study and therapy of Fanconi
PT Anaemia
XX
PS Claim 1; Page 111; 137pp; English.
XX
CC The sequences given in AAQ51429-54 represent the 5' and 3' splice
CC regions of the introns from the Fanconi Anemia Group C Complementing
CC (FACC) DNA. Three cDNA molecules which are cellular varients of a

```
CC   single cDNA, are transcribed from the FACC gene.  The three cDNAs
CC   each contain an identical open reading frame encoding the FACC
CC   protein.  FACC protein may be used for the diagnosis and study of
CC   Fanconi anemia.  The FACC gene and cDNAs may be used in gene therapy.
XX
SQ   Sequence 72 BP; 12 A; 16 C; 13 G; 31 T; 0 other;

Alignment Scores:
Pred. No.:              5.22e+03        Length:         72
Score:                  46.00           Matches:        7
Percent Similarity:     76.92%          Conservative:   3
Best Local Similarity:  53.85%          Mismatches:     3
Query Match:            10.41%          Indels:         0
DB:                     14              Gaps:           0

US-09-599-087B-5 (1-81) x AAQ51446 (1-72)

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
          ||||||||||:::  :::|||  |||  |||:::|||
Db     71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 12
AAQ36160
ID   AAQ36160 standard; DNA; 79 BP.
XX
AC   AAQ36160;
XX
DT   27-MAY-1993  (first entry)
XX
DE   Mutagenic primer COD1172.
XX
KW   Mutagenesis; site specific; nucleic acid constructs;
KW   restriction site; introduction; removal; ss.
XX
OS   Synthetic.
XX
PN   WO9301282-A.
XX
PD   21-JAN-1993.
XX
PF   01-JUL-1992;   92WO-US05573.
XX
PR   01-JUL-1991;   91US-0724237.
XX
PA   (BERL-) BERLEX LAB INC.
XX
PI   Andrews WH,  Morser MJ,  Vilander LR;
XX
DR   WPI; 1993-045488/05.
XX
PT   Site-specific mutagenesis of nucleic acid constructs - using an
PT   oligo:nucleotide which changes a nucleotide and introduces or
PT   removes a restriction site
XX
PS   Example; Page 48; 87pp; English.
XX
CC   The sequence is that of mutagenic primer COD1172 which was used as part
CC   of a method for site-specific mutagenesis of nucleic acid constructs.
CC   The primer changes a nucleotide and introduces or removes a restriction
CC   site. Using the method multiple mutations can be performed on a single
CC   target sequence simultaneously and more than one such target sequence
CC   can be included on a single construct. The method allows screening of
CC   mutants simply by restriction analysis. The primer was used in the
CC   construction of mutagenesis plasmid pTM3.0-19, it was annealed with
CC   COD1169, COD1170, COD1171 and COD1173 and inserted into the
CC   EcoRI-HindIII sites of pUC18.
XX
SQ   Sequence 79 BP; 10 A; 27 C; 32 G; 10 T; 0 other;

Alignment Scores:
Pred. No.:              5.88e+03        Length:         79
Score:                  46.00           Matches:        8
Percent Similarity:     52.94%          Conservative:   1
Best Local Similarity:  47.06%          Mismatches:     8
Query Match:            10.41%          Indels:         0
DB:                     14              Gaps:           0

US-09-599-087B-5 (1-81) x AAQ36160 (1-79)

Qy     29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValPro 45
          |||     |||:::|||  |||        |||  ||||||  |||
Db     15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCT 65

RESULT 13
AAI76614/c
ID   AAI76614 standard; DNA; 48 BP.
XX
AC   AAI76614;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:3555.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 1139; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 48 BP; 8 A; 15 C; 20 G; 5 T; 0 other;

Alignment Scores:
Pred. No.:              3.97e+03        Length:         48
Score:                  45.00           Matches:        8
Percent Similarity:     76.92%          Conservative:   2
Best Local Similarity:  61.54%          Mismatches:     3
```

```
Query Match:          10.18%      Indels:      0
DB:                   22          Gaps:        0

US-09-599-087B-5 (1-81) x AAI76614 (1-48)

Qy     67 ProGluProArgLeuTrpValValProGlyAlaLeuPro 79
          |||   |||||||||      :::||||||||||:::|||
Db     44 CCCAGTCCCCGACTCCTTGGCCTCCCAGGGGCCATCCCG 6

RESULT 14
ABN31870
ID   ABN31870 standard; DNA; 65 BP.
XX
AC   ABN31870;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:4618.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 4618; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 20 C; 13 G; 13 T; 0 other;

Alignment Scores:
Pred. No.:            6.63e+03    Length:      65
Score:                44.50       Matches:     9
Percent Similarity:   70.59%      Conservative: 3
Best Local Similarity: 52.94%     Mismatches:  2
Query Match:          10.07%      Indels:      3
DB:                   24          Gaps:        1

US-09-599-087B-5 (1-81) x ABN31870 (1-65)

Qy     43 ArgValProSerProAsnSer---------ThrAsnLeuLysGlyHisHis 56
          |||:::|||||||||   |||         |||||||||:::   |||:::
Db      1 AGAATTCCAAGCCCTCGTTCCAAAGGTGCAACCAACCTTCAGAGTCACTAT 51

RESULT 15
AAQ54486/c
ID   AAQ54486 standard; cDNA; 50 BP.
XX
AC   AAQ54486;
XX
DT   07-JUL-1994  (first entry)
XX
DE   Excitatory amino acid receptor (EAA4a) coding sequence fragment.
XX
KW   Excitatory amino acid receptor; EAA4a; screening; detection;
KW   identification; CNS; nervous disorders; antibody; oocyte; membrane
KW   protein; ion-channel; therapeutics; ss.
XX
OS   Homo sapiens.
XX
PN   EP578409-A.
XX
PD   12-JAN-1994.
XX
PF   24-JUN-1993;   93EP-0304961.
XX
PR   24-JUN-1992;   92US-0903456.
XX
PA   (ELLI/) ELLIOTT C E.
PA   (KAMB/) KAMBOJ R.
PA   (NUTT/) NUTT S L.
XX
PI   Kamboj R,  Nutt SL,  Elliott CE;
XX
DR   WPI; 1994-010248/02.
XX
PT   Excitatory aminoacid for receptors with kainate binding activity
PT   - used to test ligands for CNS receptor interaction to identify
PT   cpds. useful against CNS disorders
XX
PS   Disclosure; Figure 3b; 37pp; English.
XX
CC   The excitatory amino acid receptor (See AAR45367) can be used to
CC   screen ligands.  The ligands can be assayed for interaction by
CC   incubating them with cells, preferably oocytes, expressing the
CC   receptor or with membrane preparations from these cells, and
CC   assessing any interaction by determining receptor-ligand binding or
CC   ligand-mediated ion channel activation.  This method is useful in
CC   the identification of therapeutics useful to treat CNS disorders in
CC   humans.  Fragments of the receptor are useful for structural
CC   investigations and to raise antibodies.  This fragment of the coding
CC   sequence (AAQ54482) corresponds to bases 1971 to 2020.
XX
SQ   Sequence 50 BP; 11 A; 11 C; 15 G; 13 T; 0 other;

Alignment Scores:
Pred. No.:            5.35e+03    Length:      50
Score:                44.00       Matches:     7
Percent Similarity:   60.00%      Conservative: 2
Best Local Similarity: 46.67%     Mismatches:  6
Query Match:          9.95%       Indels:      0
DB:                   15          Gaps:        0
```

US-09-599-087B-5 (1-81) x AAQ54486 (1-50)

```
Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2
```

Search completed: May 8, 2003, 21:52:57
Job time : 210 secs

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:48:43 ; Search time 62 Seconds
(without alignments)
400.658 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 322676

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113322_7375/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_36_@runat_07052003_113322_7375 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	53	12.0	79	1	US-08-170-290A-19	Sequence 19, Appl
	2	46.5	10.5	48	2	US-08-350-260A-289	Sequence 289, App
c	3	46	10.4	72	1	US-08-441-430-22	Sequence 22, Appl
	4	45.5	10.3	69	1	US-08-554-612C-37	Sequence 37, Appl
	5	45	10.2	39	2	US-08-350-260A-233	Sequence 233, App
	6	44	10.0	42	1	US-08-642-255-23	Sequence 23, Appl
c	7	44	10.0	50	1	US-07-903-456-7	Sequence 7, Appli
c	8	44	10.0	50	1	US-07-903-456-8	Sequence 8, Appli
c	9	44	10.0	50	4	US-08-249-241-7	Sequence 7, Appli
c	10	44	10.0	50	4	US-08-249-241-8	Sequence 8, Appli
	11	44	10.0	51	2	US-08-350-260A-287	Sequence 287, App
	12	44	10.0	51	2	US-08-350-260A-290	Sequence 290, App
	13	43	9.7	72	6	5466668-7	Patent No. 5466668
	14	43	9.7	78	1	US-08-338-992B-5	Sequence 5, Appli
c	15	43	9.7	78	1	US-08-338-992B-6	Sequence 6, Appli
	16	43	9.7	78	4	US-09-010-733-5	Sequence 5, Appli
c	17	43	9.7	78	4	US-09-010-733-6	Sequence 6, Appli
	18	43	9.7	78	5	PCT-US95-09057-5	Sequence 5, Appli
c	19	43	9.7	78	5	PCT-US95-09057-6	Sequence 6, Appli
c	20	43	9.7	79	1	US-08-571-983-2	Sequence 2, Appli
c	21	43	9.7	79	5	PCT-US94-08111-2	Sequence 2, Appli
c	22	42	9.5	30	1	US-08-338-471D-4	Sequence 4, Appli
	23	42	9.5	69	1	US-07-977-284A-207	Sequence 207, App
	24	42	9.5	69	2	US-08-256-426B-207	Sequence 207, App
c	25	41.5	9.4	46	4	US-08-872-056-19	Sequence 19, Appl
	26	41.5	9.4	60	2	US-08-465-380-112	Sequence 112, App
	27	41.5	9.4	60	2	US-08-480-478-27	Sequence 27, Appl
	28	41.5	9.4	60	2	US-08-486-397-112	Sequence 112, App
	29	41.5	9.4	60	2	US-08-486-399-112	Sequence 112, App
	30	41.5	9.4	60	2	US-08-461-965-112	Sequence 112, App
	31	41.5	9.4	60	2	US-08-326-110A-27	Sequence 27, Appl
	32	41.5	9.4	60	2	US-08-634-641-112	Sequence 112, App
	33	41.5	9.4	60	3	US-09-249-471-112	Sequence 112, App
	34	41.5	9.4	60	3	US-09-249-472-112	Sequence 112, App
	35	41.5	9.4	60	3	US-09-249-451-112	Sequence 112, App
	36	41.5	9.4	60	3	US-08-809-455-112	Sequence 112, App
	37	41.5	9.4	60	3	US-09-249-461-112	Sequence 112, App
	38	41.5	9.4	60	3	US-09-249-448-112	Sequence 112, App
	39	41	9.3	30	1	US-07-969-931-27	Sequence 27, Appl
	40	41	9.3	30	1	US-07-855-417A-27	Sequence 27, Appl
	41	41	9.3	60	4	US-09-339-913B-13	Sequence 13, Appl
	42	41	9.3	60	4	US-09-339-904A-13	Sequence 13, Appl
	43	41	9.3	60	4	US-08-769-062B-13	Sequence 13, Appl
	44	41	9.3	60	4	US-09-344-002B-13	Sequence 13, Appl
	45	41	9.3	60	4	US-09-559-565C-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-170-290A-19
; Sequence 19, Application US/08170290A
; Patent No. 5702931
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Morser, Michael J.
; APPLICANT: Zielander, Laura R.
; TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James M. Heslin
; STREET: 379 Lytton Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,290A
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05573
; FILING DATE: 01-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/724,237
; FILING DATE: 01-JUL-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```
;      NAME:  Heslin, James M.
;      REGISTRATION NUMBER:  29,541
;      REFERENCE/DOCKET NUMBER:  11972-58-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  79 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-170-290A-19

Alignment Scores:
Pred. No.:                114             Length:         79
Score:                    53.00           Matches:        9
Percent Similarity:       55.00%          Conservative:   2
Best Local Similarity:    45.00%          Mismatches:     9
Query Match:              11.99%          Indels:         0
DB:                       1               Gaps:           0

US-09-599-087B-5 (1-81) x US-08-170-290A-19 (1-79)

Qy      29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      |||:::|||   |||         |||   ||||||   |||   |||:::
Db      15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCTTGCCCGAGT 74

RESULT 2
US-08-350-260A-289
; Sequence 289, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  289:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-289

Alignment Scores:
Pred. No.:                320             Length:         48
Score:                    46.50           Matches:        9
Percent Similarity:       66.67%          Conservative:   3
Best Local Similarity:    50.00%          Mismatches:     3
Query Match:              10.52%          Indels:         3
DB:                       2               Gaps:           1

US-09-599-087B-5 (1-81) x US-08-350-260A-289 (1-48)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
           ||||||   ||||||   :::   ||||||        |||:::|||:::|||
Db       3 TGGTCCCTCCGCCGAATACAGTGCTGCTGT--------CCCACGCCTGACAGT 47

RESULT 3
US-08-441-430-22/c
; Sequence 22, Application US/08441430
; Patent No. 5681942
;  GENERAL INFORMATION:
;    APPLICANT:  Buchwald, Manuel
;    APPLICANT:  Strathdee, Craig A.
;    APPLICANT:  Wevrick, Rachel
;    APPLICANT:  Mathew, Christopher George Porter
;    TITLE OF INVENTION:  Fanconi Anemia Type C Gene
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Polley, Esq.
;      ADDRESSEE:  Klarquist, Sparkman, Campbell, Leigh &
;      ADDRESSEE:  Whinston, LLP
;      STREET:  121 S.W. Salmon, Suite 1600
;      CITY:  Portland
;      STATE:  Oregon
;      COUNTRY:  U.S.A.
;      ZIP:  97204
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk, 3+-inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect 5.1/ASCII Text File
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,430
;      FILING DATE:  May 15, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. 07/876,285
;      FILING DATE:  April 29, 1992
;      APPLICATION NUMBER:  U.S. 07/918,313
;      FILING DATE:  July 21, 1992
;      APPLICATION NUMBER:  U.S. 08/003,963
;      FILING DATE:  January 15, 1993
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Richard J. Polley, Esq.
;      REGISTRATION NUMBER:  28,107
;      REFERENCE/DOCKET NUMBER:  3812-42824
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (503) 226-7391
;      TELEFAX:  (503) 228-9446
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  72 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double stranded
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-441-430-22

Alignment Scores:
Pred. No.:              642          Length:        72
Score:                  46.00        Matches:       7
Percent Similarity:     76.92%       Conservative:  3
Best Local Similarity:  53.85%       Mismatches:    3
Query Match:            10.41%       Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-441-430-22 (1-72)

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
           ||||||||||:::   :::|||   |||   |||:::|||
Db      71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 4
US-08-554-612C-37
; Sequence 37, Application US/08554612C
; Patent No. 5747660
;  GENERAL INFORMATION:
;    APPLICANT:  Orlicky, David
;    TITLE OF INVENTION:  PROSTAGLANDIN F2` RECEPTOR REGULATORY
;    TITLE OF INVENTION:  PROTEIN AND THERAPEUTIC USES
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  2200 Sand Hill Road, Suite 100
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/554,612C
;      FILING DATE:  No. 5747660ember 6, 1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  06519/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 322-5070
;      TELEFAX:  (415) 854-0875
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-554-612C-37

Alignment Scores:
Pred. No.:              691          Length:        69
Score:                  45.50        Matches:       10
Percent Similarity:     54.55%       Conservative:  2
Best Local Similarity:  45.45%       Mismatches:    7
Query Match:            10.29%       Indels:        3
DB:                     1            Gaps:          1

US-09-599-087B-5 (1-81) x US-08-554-612C-37 (1-69)

Qy      25 LysArgArgProAla---------LysAlaTrpSerGlyArgArgThrArgLeuCysCys 41
           :::||||||   |||         :::||||||      ||||||         ||||||
Db       3 AGGCGGAGAGTCGCTCCCGCCGGCCGAGCATGGGGCGCCCGGCGCCGAGGCCGCTGCTGC 62

Qy      42 HisArg 43
           |||
Db      63 TGGCGC 68

RESULT 5
US-08-350-260A-233
; Sequence 233, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  233:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-233

Alignment Scores:
Pred. No.:              357          Length:        39
Score:                  45.00        Matches:       8
Percent Similarity:     66.67%       Conservative:  0
Best Local Similarity:  66.67%       Mismatches:    4
Query Match:            10.18%       Indels:        0
DB:                     2            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-350-260A-233 (1-39)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArg 43
           ||||||  ||||||         ||||||||||||
Db       3 TGGTCCCTCCGCCGAATACCAGGCTGCTGTCATAGG 38

RESULT 6
US-08-642-255-23
; Sequence 23, Application US/08642255
; Patent No. 5773249
;  GENERAL INFORMATION:
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic"
US-08-642-255-23

Alignment Scores:
Pred. No.:              516          Length:        42
Score:                  44.00        Matches:       7
Percent Similarity:     75.00%       Conservative:  2
Best Local Similarity:  58.33%       Mismatches:    3
Query Match:            9.95%        Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-642-255-23 (1-42)

Qy      27 ArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
           ||||||  :::::::|||      ||||||||||||
Db       4 AGGCCCAGGAGGTCCTGGCGGTCCAGGCGGACCAGG 39

RESULT 7
US-07-903-456-7/c
; Sequence 7, Application US/07903456
; Patent No. 5574144
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,456
;      FILING DATE:  19920624
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/183/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-07-903-456-7

Alignment Scores:
Pred. No.:              657          Length:        50
Score:                  44.00        Matches:       7
Percent Similarity:     60.00%       Conservative:  2
Best Local Similarity:  46.67%       Mismatches:    6
Query Match:            9.95%        Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-07-903-456-7 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::     ||||||  |||  |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 8
US-07-903-456-8/c
; Sequence 8, Application US/07903456
```

```
; Patent No. 5574144
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,456
;      FILING DATE:  19920624
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/183/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-07-903-456-8

Alignment Scores:
Pred. No.:                657             Length:          50
Score:                    44.00           Matches:         7
Percent Similarity:       60.00%          Conservative:    2
Best Local Similarity:    46.67%          Mismatches:      6
Query Match:              9.95%           Indels:          0
DB:                       1               Gaps:            0

US-09-599-087B-5 (1-81) x US-07-903-456-8 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 9
US-08-249-241-7/c
; Sequence 7, Application US/08249241
; Patent No. 6376200
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K. Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,241
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,456
;      FILING DATE:  24-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/220/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-249-241-7

Alignment Scores:
Pred. No.:                657             Length:          50
Score:                    44.00           Matches:         7
Percent Similarity:       60.00%          Conservative:    2
Best Local Similarity:    46.67%          Mismatches:      6
Query Match:              9.95%           Indels:          0
DB:                       4               Gaps:            0

US-09-599-087B-5 (1-81) x US-08-249-241-7 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 10
US-08-249-241-8/c
; Sequence 8, Application US/08249241
; Patent No. 6376200
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K. Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,241
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,456
;      FILING DATE:  24-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/220/ALLE
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-249-241-8

Alignment Scores:
Pred. No.:                657             Length:          50
Score:                    44.00           Matches:         7
Percent Similarity:       60.00%          Conservative:    2
Best Local Similarity:    46.67%          Mismatches:      6
Query Match:              9.95%           Indels:          0
DB:                       4               Gaps:            0

US-09-599-087B-5 (1-81) x US-08-249-241-8 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 11
US-08-350-260A-287
; Sequence 287, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  287:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-287

Alignment Scores:
Pred. No.:                676             Length:          51
Score:                    44.00           Matches:         9
Percent Similarity:       62.50%          Conservative:    1
Best Local Similarity:    56.25%          Mismatches:      4
Query Match:              9.95%           Indels:          2
DB:                       2               Gaps:            1

US-09-599-087B-5 (1-81) x US-08-350-260A-287 (1-51)

Qy      32 TrpSerGlyArgArgThrArgLeuCys------CysHisArgValPro 45
           ||||||   ||||||   |||   |||       ||||||   :::|||
Db       3 TGGTCCCTCCGCCGAATACGAAGTTGCTCCCACTGCCATGGTCTGCCC 50

RESULT 12
US-08-350-260A-290
; Sequence 290, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  PCT/GB91/01134
;       FILING DATE:  10-JUL-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/GB92/00883
;       FILING DATE:  15-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/GB93/00605
;       FILING DATE:  24-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/150,002
;       FILING DATE:  31-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/307,619
;       FILING DATE:  16-SEP-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clough, David W
;       REGISTRATION NUMBER:  36,107
;       REFERENCE/DOCKET NUMBER:  28111/32372
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-474-6300
;   INFORMATION FOR SEQ ID NO:  290:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  51 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-350-260A-290

Alignment Scores:
Pred. No.:                676             Length:          51
Score:                    44.00           Matches:         9
Percent Similarity:       61.11%          Conservative:    2
Best Local Similarity:    50.00%          Mismatches:      5
Query Match:              9.95%           Indels:          2
DB:                       2               Gaps:            1

US-09-599-087B-5 (1-81) x US-08-350-260A-290 (1-51)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
           ||||||   ||||||         ||||||         |||:::|||:::|||
Db       3 TGGTCCCTCCGCCGAATACMNNAGTGCTGC------TGTCCCACGCCTGACAGT 50

RESULT 13
5466668-7
;Patent No. 5466668
;     APPLICANT: GLASER, CHARLES B.;MORSER, MICHAEL J.;LIGHT,
;DAVID R.
;     TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
;     NUMBER OF SEQUENCES: 57
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/155,346
;       FILING DATE: 22-NOV-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 568,456
;       FILING DATE: 15-AUG-1990
;       APPLICATION NUMBER: 506,325
;       FILING DATE: 09-APR-1990
;       APPLICATION NUMBER: 406,941
;       FILING DATE: 13-SEP-1989
;       APPLICATION NUMBER: 345,374
;       FILING DATE: 28-APR-1989
;SEQ ID NO:7:
;       LENGTH: 72
5466668-7

Alignment Scores:
Pred. No.:                1.42e+03        Length:          72
Score:                    43.00           Matches:         7
Percent Similarity:       83.33%          Conservative:    3
Best Local Similarity:    58.33%          Mismatches:      2
Query Match:              9.73%           Indels:          0
DB:                       6               Gaps:            0

US-09-599-087B-5 (1-81) x 5466668-7 (1-72)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||   ::::::|||
Db      22 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 57

RESULT 14
US-08-338-992B-5
; Sequence 5, Application US/08338992B
; Patent No. 5736524
;  GENERAL INFORMATION:
;    APPLICANT:  CONTENT, JEAN
;    APPLICANT:  HUYGEN, KRIS
;    APPLICANT:  LIU, MARGARET A.
;    APPLICANT:  MONTGOMERY, DONNA
;    APPLICANT:  ULMER, JEFFREY
;    TITLE OF INVENTION:  A POLYNUCLEOTIDE TUBERCULOSIS VACCINE
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MICHAEL D. YABLONSKY - MERCK & CO., INC.
;      STREET:  126 E. LINCOLN AVE., - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/338,992B
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YABLONSKY, MICHAEL D.
;      REGISTRATION NUMBER:  40,407
;      REFERENCE/DOCKET NUMBER:  19342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-4678
;      TELEFAX:  908-594-4720
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-338-992B-5

Alignment Scores:
Pred. No.:                1.59e+03        Length:          78
Score:                    43.00           Matches:         7
Percent Similarity:       83.33%          Conservative:    3
Best Local Similarity:    58.33%          Mismatches:      2
Query Match:              9.73%           Indels:          0
DB:                       1               Gaps:            0

US-09-599-087B-5 (1-81) x US-08-338-992B-5 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||   ::::::|||
Db      29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 15
US-08-338-992B-6/c
; Sequence 6, Application US/08338992B
; Patent No. 5736524
;  GENERAL INFORMATION:
;    APPLICANT:  CONTENT, JEAN
;    APPLICANT:  HUYGEN, KRIS
;    APPLICANT:  LIU, MARGARET A.
```

```
;    APPLICANT:  MONTGOMERY, DONNA
;    APPLICANT:  ULMER, JEFFREY
;    TITLE OF INVENTION:  A POLYNUCLEOTIDE TUBERCULOSIS VACCINE
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MICHAEL D. YABLONSKY - MERCK & CO., INC.
;      STREET:  126 E. LINCOLN AVE., - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/338,992B
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YABLONSKY, MICHAEL D.
;      REGISTRATION NUMBER:  40,407
;      REFERENCE/DOCKET NUMBER:  19342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-4678
;      TELEFAX:  908-594-4720
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-338-992B-6

Alignment Scores:
Pred. No.:              1.59e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     1               Gaps:           0

US-09-599-087B-5 (1-81) x US-08-338-992B-6 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||   |||:::|||||||||||||   ::::::|||
Db      54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19


Search completed: May  8, 2003, 22:37:37
Job time : 63 secs
```

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:49:18 ; Search time 103 Seconds
(without alignments)
977.273 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 238982

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113322_7398/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500
-MINLEN=25 -MAXLEN=80 -USER=US09599087_@CGN_1_1_93_@runat_07052003_113322_7398
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	44	10.0	50	9	US-10-126-617-7	Sequence 7, Appli
c	2	44	10.0	50	9	US-10-126-617-8	Sequence 8, Appli
	3	43	9.7	69	9	US-10-159-428-1	Sequence 1, Appli
	4	43	9.7	78	9	US-10-204-200-6	Sequence 6, Appli
c	5	43	9.7	78	9	US-10-204-200-7	Sequence 7, Appli
	6	43	9.7	78	10	US-09-950-844-5	Sequence 5, Appli
c	7	43	9.7	78	10	US-09-950-844-6	Sequence 6, Appli
	8	43	9.7	78	10	US-09-835-699-5	Sequence 5, Appli
c	9	43	9.7	78	10	US-09-835-699-6	Sequence 6, Appli
c	10	42.5	9.6	59	9	US-10-006-593-89	Sequence 89, Appl
	11	42	9.5	50	9	US-10-137-788-6	Sequence 6, Appli
c	12	41.5	9.4	46	10	US-09-827-864-19	Sequence 19, Appl
c	13	41.5	9.4	54	10	US-09-294-093B-4383	Sequence 4383, Ap
	14	41.5	9.4	60	9	US-09-902-941-1979	Sequence 1979, Ap
	15	41.5	9.4	60	9	US-10-017-754-1979	Sequence 1979, Ap
	16	41	9.3	60	9	US-09-954-692-13	Sequence 13, Appl
	17	41	9.3	60	10	US-09-559-671A-13	Sequence 13, Appl
	18	41	9.3	80	9	US-09-440-829-17	Sequence 17, Appl
c	19	41	9.3	80	10	US-09-157-748-43	Sequence 43, Appl
c	20	40.5	9.2	75	10	US-09-864-761-28813	Sequence 28813, A
c	21	40	9.0	31	9	US-09-773-599-16	Sequence 16, Appl
c	22	40	9.0	31	10	US-09-801-274-1511	Sequence 1511, Ap
c	23	40	9.0	31	10	US-09-801-274-1512	Sequence 1512, Ap
	24	40	9.0	46	9	US-09-940-244-227	Sequence 227, App
	25	40	9.0	48	9	US-09-877-478-5348	Sequence 5348, Ap
c	26	40	9.0	59	9	US-10-025-380-976	Sequence 976, App
c	27	40	9.0	59	10	US-09-922-217-976	Sequence 976, App
c	28	40	9.0	59	10	US-09-833-263-976	Sequence 976, App
c	29	40	9.0	59	12	US-10-013-737-15	Sequence 15, Appl
	30	40	9.0	60	9	US-09-902-941-1980	Sequence 1980, Ap
	31	40	9.0	60	9	US-10-017-754-1980	Sequence 1980, Ap
	32	40	9.0	63	10	US-09-923-876-1639	Sequence 1639, Ap
c	33	40	9.0	64	9	US-10-072-036-7	Sequence 7, Appli
c	34	40	9.0	75	9	US-09-979-999-15	Sequence 15, Appl
	35	40	9.0	75	9	US-10-045-465-2	Sequence 2, Appli
	36	40	9.0	78	10	US-09-864-761-21687	Sequence 21687, A
	37	39.5	8.9	78	9	US-09-861-257-47	Sequence 47, Appl
	38	39	8.8	25	9	US-10-215-112-13778	Sequence 13778, A
	39	39	8.8	28	9	US-10-156-634A-13	Sequence 13, Appl
	40	39	8.8	36	9	US-09-345-373-5	Sequence 5, Appli
	41	39	8.8	37	10	US-09-263-959-815	Sequence 815, App
	42	39	8.8	40	10	US-09-263-959-458	Sequence 458, App
	43	39	8.8	64	9	US-10-085-906-168	Sequence 168, App
c	44	39	8.8	65	10	US-09-747-003-12	Sequence 12, Appl
	45	39	8.8	69	9	US-10-076-816-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-10-126-617-7/c
; Sequence 7, Application US/10126617
; Patent No. US20020168723A1
; GENERAL INFORMATION:
; APPLICANT: KAMBOJ, RAJENDER
; APPLICANT: ELLIOTT, CANDACE E.
; APPLICANT: NUTT, STEPHEN L.
; TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF THE EAA4 FAMILY
; FILE REFERENCE: 016777/0471
; CURRENT APPLICATION NUMBER: US/10/126,617
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 08/249,241
; PRIOR FILING DATE: 1994-05-25
; PRIOR APPLICATION NUMBER: 07/903,456
; PRIOR FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-617-7

Alignment Scores:

Pred. No.:	566	Length:	50
Score:	44.00	Matches:	7
Percent Similarity:	60.00%	Conservative:	2

```
Best Local Similarity:  46.67%          Mismatches:     6
Query Match:            9.95%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-126-617-7 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||       :::       ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 2
US-10-126-617-8/c
; Sequence 8, Application US/10126617
; Patent No. US20020168723A1
; GENERAL INFORMATION:
;  APPLICANT: KAMBOJ, RAJENDER
;  APPLICANT:  ELLIOTT, CANDACE E.
;  APPLICANT:  NUTT, STEPHEN L.
;  TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF THE EAA4 FAMILY
;  FILE REFERENCE: 016777/0471
;  CURRENT APPLICATION NUMBER: US/10/126,617
;  CURRENT FILING DATE:  2002-04-22
;  PRIOR APPLICATION NUMBER: 08/249,241
;  PRIOR FILING DATE: 1994-05-25
;  PRIOR APPLICATION NUMBER: 07/903,456
;  PRIOR FILING DATE: 1992-06-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-126-617-8

Alignment Scores:
Pred. No.:              566             Length:         50
Score:                  44.00           Matches:        7
Percent Similarity:     60.00%          Conservative:   2
Best Local Similarity:  46.67%          Mismatches:     6
Query Match:            9.95%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-126-617-8 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||       :::       ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 3
US-10-159-428-1
; Sequence 1, Application US/10159428
; Publication No. US20030082579A1
; GENERAL INFORMATION:
;  APPLICANT: Felgner, Philip L
;  APPLICANT:  Doolan, Denise L
;  TITLE OF INVENTION: PROTEIN ARRAYS AND METHODS AND SYSTEMS
;  TITLE OF INVENTION:  FOR PRODUCING THE SAME
;  FILE REFERENCE: GTSYS.008A
;  CURRENT APPLICATION NUMBER: US/10/159,428
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: 60/294,739
;  PRIOR FILING DATE: 2001-05-30
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-159-428-1

Alignment Scores:
Pred. No.:              1.13e+03        Length:         69
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-159-428-1 (1-69)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::||||||||||||   ::::::|||
Db      22 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 57

RESULT 4
US-10-204-200-6
; Sequence 6, Application US/10204200
; Publication No. US20030036108A1
; GENERAL INFORMATION:
;  APPLICANT: Sullivan, Kathleen A.
;  APPLICANT:  Benincasa, Diana
;  APPLICANT:  Cascieri, Margaret A.
;  APPLICANT:  Mitnaul, Lyndon J.
;  APPLICANT:  Shiao, Lin-Lin
;  APPLICANT:  Tota, Michael A.
;  TITLE OF INVENTION: CELL FUSION ASSAYS USING FLUORESCENCE
;  TITLE OF INVENTION:  RESONANCE ENERGY TRANSFER
;  FILE REFERENCE: 20618P
;  CURRENT APPLICATION NUMBER: US/10/204,200
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: PCT/US01/04677
;  PRIOR FILING DATE: 2001-02-13
;  PRIOR APPLICATION NUMBER: 60/183,309
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Human
US-10-204-200-6

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-204-200-6 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::||||||||||||   ::::::|||
Db      29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 5
US-10-204-200-7/c
; Sequence 7, Application US/10204200
; Publication No. US20030036108A1
; GENERAL INFORMATION:
;  APPLICANT: Sullivan, Kathleen A.
;  APPLICANT:  Benincasa, Diana
;  APPLICANT:  Cascieri, Margaret A.
;  APPLICANT:  Mitnaul, Lyndon J.
;  APPLICANT:  Shiao, Lin-Lin
;  APPLICANT:  Tota, Michael A.
;  TITLE OF INVENTION: CELL FUSION ASSAYS USING FLUORESCENCE
;  TITLE OF INVENTION:  RESONANCE ENERGY TRANSFER
;  FILE REFERENCE: 20618P
;  CURRENT APPLICATION NUMBER: US/10/204,200
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: PCT/US01/04677
;  PRIOR FILING DATE: 2001-02-13
;  PRIOR APPLICATION NUMBER: 60/183,309
;  PRIOR FILING DATE: 2000-02-17
```

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Human
US-10-204-200-7

Alignment Scores:

Pred. No.:	1.32e+03	Length:	78
Score:	43.00	Matches:	7
Percent Similarity:	83.33%	Conservative:	3
Best Local Similarity:	58.33%	Mismatches:	2
Query Match:	9.73%	Indels:	0
DB:	9	Gaps:	0

US-09-599-087B-5 (1-81) x US-10-204-200-7 (1-78)

```
Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::|||||||||||||   ::::::|||
Db     54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19
```

RESULT 6
US-09-950-844-5
; Sequence 5, Application US/09950844
; Patent No. US20020045594A1
; GENERAL INFORMATION:
; APPLICANT: Volkin, David B.
; APPLICANT: Evans, Robert K.
; APPLICANT: Ulmer, Jeffrey B.
; APPLICANT: Caulfield, Michael J.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULATIONS
; FILE REFERENCE: 19907YIACA
; CURRENT APPLICATION NUMBER: US/09/950,844
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/112,655
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 09/023,834
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/038,194
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-950-844-5

Alignment Scores:

Pred. No.:	1.32e+03	Length:	78
Score:	43.00	Matches:	7
Percent Similarity:	83.33%	Conservative:	3
Best Local Similarity:	58.33%	Mismatches:	2
Query Match:	9.73%	Indels:	0
DB:	10	Gaps:	0

US-09-599-087B-5 (1-81) x US-09-950-844-5 (1-78)

```
Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::|||||||||||||   ::::::|||
Db     29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64
```

RESULT 7
US-09-950-844-6/c
; Sequence 6, Application US/09950844
; Patent No. US20020045594A1
; GENERAL INFORMATION:
; APPLICANT: Volkin, David B.
; APPLICANT: Evans, Robert K.
; APPLICANT: Ulmer, Jeffrey B.
; APPLICANT: Caulfield, Michael J.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULATIONS
; FILE REFERENCE: 19907YIACA
; CURRENT APPLICATION NUMBER: US/09/950,844
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/112,655
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 09/023,834
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/038,194
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-950-844-6

Alignment Scores:

Pred. No.:	1.32e+03	Length:	78
Score:	43.00	Matches:	7
Percent Similarity:	83.33%	Conservative:	3
Best Local Similarity:	58.33%	Mismatches:	2
Query Match:	9.73%	Indels:	0
DB:	10	Gaps:	0

US-09-599-087B-5 (1-81) x US-09-950-844-6 (1-78)

```
Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::|||||||||||||   ::::::|||
Db     54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19
```

RESULT 8
US-09-835-699-5
; Sequence 5, Application US/09835699
; Patent No. US20020147167A1
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Marcy E.
; Keys, Robert D.
; Lewis, John A.
; Liu, Margaret A.
; McClements, William L.
; TITLE OF INVENTION: A POLYNUCLEOTIDE HERPES VIRUS VACCINE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/835,699
; FILING DATE: 17-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/17262
; FILING DATE: 26-SEP-1997
; APPLICATION NUMBER: US 08/720,758
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19258CC

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 732-594-3905
;              TELEFAX: 732-594-4720
;              TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 5:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 78 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Other
;         FEATURE:
;              NAME/KEY:
;              LOCATION:  1...78
;              OTHER INFORMATION:  nucleic acid
;         SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-835-699-5

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-835-699-5 (1-78)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::|||||||||||||   ::::::|||
Db     29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 9
US-09-835-699-6/c
; Sequence 6, Application US/09835699
; Patent No. US20020147167A1
;   GENERAL INFORMATION:
;         APPLICANT: Armstrong, Marcy E.
;                    Keys, Robert D.
;                    Lewis, John A.
;                    Liu, Margaret A.
;                    McClements, William L.
;         TITLE OF INVENTION: A POLYNUCLEOTIDE HERPES VIRUS
;                             VACCINE
;         NUMBER OF SEQUENCES: 15
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Merck & Co., Inc.
;              STREET: P.O. Box 2000, 126 E. Lincoln Ave.
;              CITY: Rahway
;              STATE: NJ
;              COUNTRY: USA
;              ZIP: 07065-0907
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Diskette
;              COMPUTER: IBM Compatible
;              OPERATING SYSTEM: Windows
;              SOFTWARE: FastSEQ for Windows Version 2.0b
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/835,699
;              FILING DATE: 17-Aug-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: PCT/US97/17262
;              FILING DATE: 26-SEP-1997
;              APPLICATION NUMBER: US 08/720,758
;              FILING DATE: 01-OCT-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Hand, J. Mark
;              REGISTRATION NUMBER: 36,545
;              REFERENCE/DOCKET NUMBER: 19258CC
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 732-594-3905
;              TELEFAX: 732-594-4720
;              TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 78 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Other
;         FEATURE:
;              NAME/KEY:
;              LOCATION:  1...78
;              OTHER INFORMATION:  nucleic acid
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-835-699-6

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-835-699-6 (1-78)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::|||||||||||||   ::::::|||
Db     54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19

RESULT 10
US-10-006-593-89/c
; Sequence 89, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
;  FILE REFERENCE: 1087-2
;  CURRENT APPLICATION NUMBER: US/10/006,593
;  CURRENT FILING DATE:  2001-12-05
;  PRIOR APPLICATION NUMBER: US 60/251,448
;  PRIOR FILING DATE: 2000-12-05
;  PRIOR APPLICATION NUMBER: US 60/288,889
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/294,068
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-006-593-89

Alignment Scores:
Pred. No.:              1.07e+03        Length:         59
Score:                  42.50           Matches:        9
Percent Similarity:     57.89%          Conservative:   2
Best Local Similarity:  47.37%          Mismatches:     7
Query Match:            9.62%           Indels:         1
DB:                     9               Gaps:           1

US-09-599-087B-5 (1-81) x US-10-006-593-89 (1-59)

Qy     15 LeuCysPheSerIlePheSerThrGluGlyLysArgArgPro---AlaLysAlaTrp 32
          |||   ||||||:::|||   |||      |||   :::|||   |||      |||
Db     59 CTGCTATTTTCATTATTCAAAACCCGATCGAAGGCCCAACCCTGCGCCAGTGGCTGG 3

RESULT 11
US-10-137-788-6
```

```
; Sequence 6, Application US/10137788
; Publication No. US20030039636A1
; GENERAL INFORMATION:
;  APPLICANT: Leboulch, Philippe
;  APPLICANT:  Westerman, Karen
;  TITLE OF INVENTION: NOVEL SELF-INACTIVATING (SIN) LENTIVIRAL VECTORS
;  FILE REFERENCE: IOI-023
;  CURRENT APPLICATION NUMBER: US/10/137,788
;  CURRENT FILING DATE:  2002-05-01
;  PRIOR APPLICATION NUMBER: US 60/288,042
;  PRIOR FILING DATE: 2001-05-01
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
US-10-137-788-6

Alignment Scores:
Pred. No.:              1.01e+03        Length:         50
Score:                  42.00           Matches:        7
Percent Similarity:     73.33%          Conservative:   4
Best Local Similarity:  46.67%          Mismatches:     4
Query Match:            9.50%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-137-788-6 (1-50)

Qy      4 LeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSer 18
          |||   ||||||:::   :::|||:::|||:::   |||   |||
Db      4 CTTGCCTTGAGTGCTTCAATGTGTGTGTTGGTTTTTTGTGTGTCG 48

RESULT 12
US-09-827-864-19/c
; Sequence 19, Application US/09827864
; Patent No. US20020009458A1
;   GENERAL INFORMATION:
;        APPLICANT: COLAU, DIDIER
;                   ROOS, JOEL
;        TITLE OF INVENTION: RECOMBINANT DNA SEQUENCES, MOLECULES,
;                            VECTORS AND VACCINES FOR FELINE
;                            CALICIVIRUS DISEASE AND METHODS FOR
;                            PRODUCING AND USING SAME
;        NUMBER OF SEQUENCES: 24
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MCDERMOTT, WILL & EMERY
;             STREET: 1850 K STREET, N.W., SUITE 500
;             CITY: WASHINGTON
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20006
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/827,864
;             FILING DATE: 06-Apr-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: GADIANO, WILHLEM F
;             REGISTRATION NUMBER: 37,136
;             REFERENCE/DOCKET NUMBER: 37712-213
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 778-8373
;             TELEFAX: (202) 778-8335
;   INFORMATION FOR SEQ ID NO: 19:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 46 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: nucleic acid
;          SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-827-864-19

Alignment Scores:
Pred. No.:              1.05e+03        Length:         46
Score:                  41.50           Matches:        8
Percent Similarity:     64.29%          Conservative:   1
Best Local Similarity:  57.14%          Mismatches:     4
Query Match:            9.39%           Indels:         1
DB:                     10              Gaps:           1

US-09-599-087B-5 (1-81) x US-09-827-864-19 (1-46)

Qy      60 CysLysProCysLysLeuGluProGluProArgLeuTrpVal 73
             |||   ||||||||||||||||      |||      ||:::
Db      42 TGTCACCCTTGCAAGCTGGAATGACGTCCTTCT---TGGTTA 4

RESULT 13
US-09-294-093B-4383/c
; Sequence 4383, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:      PERL Program
; SEQ ID NO 4383
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700353949H1
;   NAME/KEY: unsure
;   LOCATION: 34
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4383

Alignment Scores:
Pred. No.:              1.28e+03        Length:         54
Score:                  41.50           Matches:        8
Percent Similarity:     69.23%          Conservative:   1
Best Local Similarity:  61.54%          Mismatches:     1
Query Match:            9.39%           Indels:         3
DB:                     10              Gaps:           1

US-09-599-087B-5 (1-81) x US-09-294-093B-4383 (1-54)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSer 46
           |||||||||         |||||||||   :::|||||||
Db      52 GGACGGAGG---------TGCTGCCATNTATTACCAAGT 23

RESULT 14
US-09-902-941-1979
; Sequence 1979, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
```

; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1979
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1979

Alignment Scores:

Pred. No.:	1.46e+03	Length:	60
Score:	41.50	Matches:	9
Percent Similarity:	61.11%	Conservative:	2
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	9.39%	Indels:	1
DB:	9	Gaps:	1

US-09-599-087B-5 (1-81) x US-09-902-941-1979 (1-60)

```
Qy      34 GlyArgArgThrArgLeu---CysCysHisArgValProSerProAsnSerThr 50
           |||   :::   :::|||   ||||||         |||||||||   ||||||
Db       1 GGGGCTCAGGCGAAGCTGGGGTGCTGTTGGGGGTATCCGAGTCCCAGAAGCACC 54
```

RESULT 15
US-10-017-754-1979
; Sequence 1979, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1979
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1979

Alignment Scores:

Pred. No.:	1.46e+03	Length:	60
Score:	41.50	Matches:	9
Percent Similarity:	61.11%	Conservative:	2
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	9.39%	Indels:	1
DB:	9	Gaps:	1

US-09-599-087B-5 (1-81) x US-10-017-754-1979 (1-60)

```
Qy      34 GlyArgArgThrArgLeu---CysCysHisArgValProSerProAsnSerThr 50
           |||   :::   :::|||   ||||||         |||||||||   ||||||
Db       1 GGGGCTCAGGCGAAGCTGGGGTGCTGTTGGGGGTATCCGAGTCCCAGAAGCACC 54
```

Search completed: May 8, 2003, 22:39:34
Job time : 104 secs

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:41:38 ; Search time 35 Seconds
(without alignments)
308.380 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 219634

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	326	73.8	57	23	AAE16482	Human Secreted epi
2	228.5	51.7	78	23	AAE16479	Mouse Secreted epi
3	183	41.4	77	23	AAE16483	Rat Secreted epith
4	169	38.2	30	23	AAE16485	Human Secreted epi
5	163	36.9	28	23	AAE16486	Human Secreted epi
6	132.5	30.0	54	23	AAE16480	Mouse Secreted epi
7	62	14.0	78	22	AAU45851	Propionibacterium
8	61	13.8	75	22	AAU66859	Propionibacterium
9	61	13.8	75	22	AAU54516	Propionibacterium
10	59.5	13.5	73	21	AAG00379	Human secreted pro
11	58.5	13.2	77	20	AAY36266	Human secreted pro
12	58	13.1	53	19	AAW75178	Human secreted pro
13	58	13.1	62	22	AAO05651	Human polypeptide
14	58	13.1	63	21	AAB44937	Human secreted pro
15	58	13.1	63	22	AAB60725	Human secreted pro
16	57.5	13.0	74	21	AAG12200	Arabidopsis thalia
17	57.5	13.0	76	21	AAG12199	Arabidopsis thalia
18	57	12.9	78	22	ABG08705	Novel human diagno
19	56	12.7	58	22	AAB60722	Human secreted pro
20	56	12.7	61	22	AAO12472	Human polypeptide
21	56	12.7	65	22	AAU43919	Propionibacterium
22	55.5	12.6	53	20	AAY27661	Human secreted pro
23	55.5	12.6	64	22	AAU40623	Propionibacterium
24	55	12.4	50	22	AAU61037	Propionibacterium
25	55	12.4	50	22	AAU64737	Propionibacterium
26	55	12.4	80	20	AAY36650	Fragment of human
27	54.5	12.3	76	22	AAU48701	Propionibacterium
28	54	12.2	48	14	AAR34681	T-PA signal peptid
29	54	12.2	57	22	AAU62049	Propionibacterium
30	54	12.2	68	22	AAU40407	Propionibacterium
31	54	12.2	70	21	AAB63096	Human secreted pro
32	53.5	12.1	44	22	ABB43990	Peptide #11496 enc
33	53.5	12.1	44	22	AAM64993	Human brain expres
34	53.5	12.1	44	22	AAM21629	Peptide #8063 enco
35	53.5	12.1	44	22	AAM37931	Peptide #11968 enc
36	53.5	12.1	44	23	ABG46742	Human peptide enco
37	53.5	12.1	52	22	AAU56061	Propionibacterium
38	53.5	12.1	53	22	ABG25839	Novel human diagno
39	53.5	12.1	62	22	AAM85736	Human immune/haema
40	53.5	12.1	71	22	AAU60257	Propionibacterium
41	53.5	12.1	71	23	ABP07526	Human ORFX protein
42	53.5	12.1	74	22	AAO12569	Human polypeptide
43	53	12.0	51	22	AAU65879	Propionibacterium
44	53	12.0	61	22	AAU33442	Enterococcus faeca
45	53	12.0	61	22	AAU54661	Propionibacterium

ALIGNMENTS

RESULT 1
AAE16482
ID AAE16482 standard; peptide; 57 AA.
XX
AC AAE16482;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) protein fragment.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human.
XX
OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX

```
DR   WPI; 2002-122281/16.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 14; Page 122; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   protein fragment.
XX
SQ   Sequence   57 AA;

  Query Match            73.8%;  Score 326;  DB 23;  Length 57;
  Best Local Similarity  100.0%;  Pred. No. 1e-30;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 81
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 57


RESULT 2
AAE16479
ID   AAE16479 standard; Protein; 78 AA.
XX
AC   AAE16479;
XX
DT   09-APR-2002  (first entry)
XX
DE   Mouse Secreted epithelial colon stromal-1 (Secs-1) protein.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   mouse.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= Signal_peptide
FT   Protein         25..78
FT                   /label= Mature_mouse_Secs-1_protein
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   N-PSDB; AAD27024.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 13; Fig 3; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is mouse Secs-1
CC   protein.
XX
SQ   Sequence   78 AA;

  Query Match            51.7%;  Score 228.5;  DB 23;  Length 78;
  Best Local Similarity  55.6%;  Pred. No. 3.5e-19;
  Matches   45;  Conservative   10;  Mismatches   23;  Indels    3;  Gaps    1;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKCHHVRLC 60
           |||| || |||:||||| |||:||:| |||:   ||   |||  |    |  ||:| | |
Db       1 MRLLALSGLLCMLLLCFCIFSSEGRRHPAKSLKLRR---CCHLSPRSKLTTWKGNHTRPC 57

Qy      61 KPCKLEPEPRLWVVPGALPQV 81
           : |: :   : |||||||||:
Db      58 RLCRNKLPVKSWVVPGALPQI 78


RESULT 3
AAE16483
ID   AAE16483 standard; Protein; 77 AA.
XX
AC   AAE16483;
XX
DT   09-APR-2002  (first entry)
XX
DE   Rat Secreted epithelial colon stromal-1 (Secs-1) protein.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   rat.
XX
OS   Rattus norvegicus.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
```

```
XX
DR   WPI; 2002-122281/16.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 3; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is rat Secs-1
CC   protein.
XX
SQ   Sequence   77 AA;

  Query Match             41.4%;  Score 183;  DB 23;  Length 77;
  Best Local Similarity   46.9%;  Pred. No. 7.2e-14;
  Matches   38;  Conservative   11;  Mismatches   28;  Indels    4;  Gaps    2;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLC 60
           |||| || |  :| ||  : |:||::||||     : |  ||  |       ||:| | |
Db       1 MRLLTLSGLFFMLFLCLCVLSSEGRKRPAKF---PKLRPRCHLSPRSKPITWKGNHTRPC 57

Qy      61 KPCKLEPEPRLWVVPGALPQV 81
             :||: : |   |||||||||:
Db      58 RPCR-KLESNSWVVPGALPQI 77


RESULT 4
AAE16485
ID   AAE16485 standard; peptide; 30 AA.
XX
AC   AAE16485;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human Secreted epithelial colon stromal-1 (Secs-1) peptide #2.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   N-PSDB; AAD27026.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 4D; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   peptide encoded by second exon.
XX
SQ   Sequence   30 AA;

  Query Match             38.2%;  Score 169;  DB 23;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      24 GKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 53
           ||||||||||||||||||||||||||||||
Db       1 GKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 30


RESULT 5
AAE16486
ID   AAE16486 standard; peptide; 28 AA.
XX
AC   AAE16486;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human Secreted epithelial colon stromal-1 (Secs-1) peptide #3.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   N-PSDB; AAD27026.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
```

```
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 4F; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   peptide encoded by third exon.
XX
SQ   Sequence   28 AA;

  Query Match             36.9%;  Score 163;  DB 23;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-12;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      54 GHHVRLCKPCKLEPEPRLWVVPGALPQV 81
           ||||||||||||||||||||||||||||
Db       1 GHHVRLCKPCKLEPEPRLWVVPGALPQV 28


RESULT 6
AAE16480
ID   AAE16480 standard; peptide; 54 AA.
XX
AC   AAE16480;
XX
DT   09-APR-2002  (first entry)
XX
DE   Mouse Secreted epithelial colon stromal-1 (Secs-1) protein fragment.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   mouse.
XX
OS   Mus musculus.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 14; Page 120; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is mouse Secs-1
CC   protein fragment.
XX
SQ   Sequence   54 AA;

  Query Match             30.0%;  Score 132.5;  DB 23;  Length 54;
  Best Local Similarity   45.6%;  Pred. No. 3.9e-08;
  Matches   26;  Conservative    8;  Mismatches   20;  Indels    3;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 81
           :| |||:   ||   ||| |    |  ||:| | |: |: :   : |||||||||:
Db       1 RRHPAKSLKLRR---CCHLSPRSKLTTWKGNHTRPCRLCRNKLPVKSWVVPGALPQI 54


RESULT 7
AAU45851
ID   AAU45851 standard; Protein; 78 AA.
XX
AC   AAU45851;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #6747.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
DR   N-PSDB; AAS59528.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Example 1; SEQ ID No 7046; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
```

CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 78 AA;

Query Match 14.0%; Score 62; DB 22; Length 78;
Best Local Similarity 78.6%; Pred. No. 9.9;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 25 KRRPAKAWSGRRTR 38
:||| ||||||| |
Db 50 RRRPRKAWSGRRRR 63

RESULT 8
AAU66859
ID AAU66859 standard; Protein; 75 AA.
XX
AC AAU66859;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #27755.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59761.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 28054; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 75 AA;

Query Match 13.8%; Score 61; DB 22; Length 75;
Best Local Similarity 34.3%; Pred. No. 12;
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

Qy 41 CHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVP 75
| | | :: : || : ||:|| | | |
Db 32 CRSVASSSTASAKGWSLALCQPCLRPLTPVSWCRP 66

RESULT 9
AAU54516
ID AAU54516 standard; Protein; 75 AA.
XX
AC AAU54516;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #15412.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59565.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 15711; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

```
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   75 AA;

  Query Match             13.8%;  Score 61;  DB 22;  Length 75;
  Best Local Similarity   34.3%;  Pred. No. 12;
  Matches   12;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      41 CHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVP 75
           |  | | :: : || -: ||:||     |  |  |
Db      32 CRSVASSSTASAKGWSLALCQPCLRPLTPVSWCRP 66


RESULT 10
AAG00379
ID   AAG00379 standard; Protein; 73 AA.
XX
AC   AAG00379;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4460.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC00385.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 4460; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   73 AA;

  Query Match             13.5%;  Score 59.5;  DB 21;  Length 73;
  Best Local Similarity   26.8%;  Pred. No. 18;
  Matches   22;  Conservative    7;  Mismatches   16;  Indels   37;  Gaps    4;

Qy       2 RLLVLS-------SLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKG 54
           :||||         | | :| || :| | | :  |:|
Db      12 QLLVLHWCVLFCFVLFCFVLFCFVLFET-GSQSVAQA---------------------- 47

Qy      55 HHVRLCKPCKLEPEPRLWVVPG 76
             |: |  | |:| |     ||
Db      48 -GVQSCDLCSLQPPP-----PG 63


RESULT 11
AAY36266
ID   AAY36266 standard; Protein; 77 AA.
XX
AC   AAY36266;
XX
DT   17-SEP-1999  (first entry)
XX
DE   Human secreted protein encoded by gene 43.
XX
KW   Human; secreted protein; cancer; tumour; developmental abnormality;
KW   foetal deficiency; blood disorder; immune system disorder; inflammation;
KW   autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW   schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW   atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW   digestive disorder; endocrine disorder; infection; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO9931117-A1.
XX
PD   24-JUN-1999.
XX
PF   17-DEC-1998;   98WO-US27059.
XX
PR   19-DEC-1997;   97US-0068369.
PR   18-DEC-1997;   97US-0068006.
PR   18-DEC-1997;   97US-0068007.
PR   18-DEC-1997;   97US-0068008.
PR   18-DEC-1997;   97US-0068053.
PR   18-DEC-1997;   97US-0068054.
PR   18-DEC-1997;   97US-0068057.
PR   18-DEC-1997;   97US-0068064.
PR   18-DEC-1997;   97US-0070923.
PR   19-DEC-1997;   97US-0068169.
PR   19-DEC-1997;   97US-0068365.
PR   19-DEC-1997;   97US-0068367.
PR   19-DEC-1997;   97US-0068368.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Carter KC,  Duan RD,  Feng P,  Ferrie AM,  Florence C;
PI   Florence K,  Greene JM,  Janat F,  Kyaw H,  Moore PA;
PI   Ni J,  Rosen CA,  Ruben SM,  Shi Y,  Soppet DR,  Wei Y;
PI   Yu G;
XX
DR   WPI; 1999-418749/35.
DR   N-PSDB; AAX97958.
XX
PT   New isolated human genes encoding secreted polypeptides
XX
PS   Claim 11; Page 368; 537pp; English.
XX
CC   AAX97916 to AAX98029 represent 110 isolated human secreted protein
```

CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by
CC the 110 human genes. The genes and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new genes.
CC Specific uses are described for each of the 110 genes, based on which
CC tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours, developmental
CC abnormalities and foetal deficiencies, blood disorders, diseases of the
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
CC disorders, kidney disorders, digestive/endocrine disorders, infections
CC and AIDS. The polypeptides are also useful for identifying their binding
CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
CC used in the exemplification of the present invention.
XX
SQ Sequence 77 AA;

Query Match 13.2%; Score 58.5; DB 20; Length 77;
Best Local Similarity 32.8%; Pred. No. 25;
Matches 19; Conservative 5; Mismatches 17; Indels 17; Gaps 2;

```
Qy       4 LVLSSLLCILLLCF--------------SIFSTEGKR--RPAKAWSGRRTRLCCHRV 44
           ||| |:  |||:                 |:  | |:|  || :   |   || | ||
Db      11 LVLGSVFMILLIIVYWDSAAPRTSTCTRPSLGRTRGRRCPRPGRTGQGAHGRLRCRRV 68
```

RESULT 12
AAW75178
ID AAW75178 standard; Protein; 53 AA.
XX
AC AAW75178;
XX
DT 28-JAN-1999 (first entry)
XX
DE Human secreted protein encoded by gene 69 clone HPEBD70.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9839446-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04492.
XX
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043315.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.

```
PR   22-AUG-1997;   97US-0056911.
PR   05-SEP-1997;   97US-0057650.
PR   05-SEP-1997;   97US-0057761.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Bednarik DP,  Brewer LA,  Carter KC,  Duan R,  Ebner R,  Endress GA;
PI   Feng P,  Ferrie AM,  Fischer CL,  Graves KA,  Greene JM,  Hu JS;
PI   Kyaw H,  Lafleur DW,  Li Y,  Moore PA,  Ni J,  Olsen HS,  Rosen CA;
PI   Ruben SM,  Shi Y,  Soppet DR,  Young PE,  Yu GL,  Zeng Z;
XX
DR   WPI; 1998-609887/51.
DR   N-PSDB; AAV34275.
XX
PT   New isolated human genes and the secreted polypeptides they encode
PT   - useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Claim 1; Page 352; 447pp; English.
XX
CC   This sequence represents a secreted human protein encoded by the gene
CC   clone detailed in the descriptor line.
CC   The gene can be used to generate fusion proteins by linking to the gene
CC   to a human immunoglobulin Fc portion (e.g. AAV34145) for increasing the
CC   stability of the fused protein as compared to the human protein only.
CC   The invention relates to 70 novel genes and their fragments (nucleic
CC   acid sequences: AAV34154-V34276; amino acid sequences AAW75057-W75179)
CC   which are useful for preventing, treating or ameliorating  medical
CC   conditions e.g. by protein or gene therapy.  Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 70
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAV34154 for described uses).
XX
SQ   Sequence   53 AA;

  Query Match             13.1%;  Score 58;  DB 19;  Length 53;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches   12;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 LLVLSSLLCILLLCFSIFSTEGKR 26
           ||:|:||  : | |: :| |:|||
Db      19 LLLLASLAPLWLYCWRMFPTKGKR 42


RESULT 13
AAO05651
ID   AAO05651 standard; Protein; 62 AA.
XX
AC   AAO05651;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 19543.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI85582.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 20; SEQ ID NO 19543; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   62 AA;

  Query Match             13.1%;  Score 58;  DB 22;  Length 62;
  Best Local Similarity   28.6%;  Pred. No. 23;
  Matches   18;  Conservative    5;  Mismatches   30;  Indels   10;  Gaps    3;

Qy      10 LCILLLCFSIFSTEGKR--RPAKAWSGRRTRLCCHRVPSPNSTNLK-GHHVRLCKPCKLE 66
           || || |   | || :   :   |         |  | |    ||   |: | :
Db       2 LCCLLFCCCYFETESRSVVQAGVQWHD-------HSSPQPGPPRLKSSSHLSLWRHDHSS 54

Qy      67 PEP 69
           |:|
Db      55 PQP 57


RESULT 14
AAB44937
ID   AAB44937 standard; Protein; 63 AA.
XX
AC   AAB44937;
XX
DT   12-FEB-2001  (first entry)
XX
DE   Human secreted protein encoded by gene 21.
XX
KW   Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
KW   antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW   nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW   ophthalmological; gene therapy; treatment; autoimmune disease; infection;
KW   cardiovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; epithelial cell proliferation; skin aging; mental state;
KW   transplantation; metabolism modulation.
XX
OS   Homo sapiens.
XX
PN   WO200055200-A1.
XX
PD   21-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06042.
XX
PR   12-MAR-1999;   99US-0124143.
PR   03-DEC-1999;   99US-0168663.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
```

PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-656008/63.
DR N-PSDB; AAC79918.
XX
PT Isolated human secretory proteins, nucleic acids encoding them and
PT antibodies directed against them, useful for diagnosing and treating
PT disorders related to the proteins such as cancer, Alzheimer's disease
PT and Parkinsons -
XX
PS Claim 11; Page 388; 453pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) and its
CC encoding nucleic acid molecule (II) which have immunosuppressive,
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant,
CC vasotropic, cerebroprotective, nootropic, neuroprotective, antibacterial,
CC virucide, fungicide and ophthalmological activity and which can be used
CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a
CC pathological condition or susceptibility to a pathological condition. The
CC antibodies to (II) can also be used in alleviating symptoms associated
CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays
CC or enzyme linked immunosorbent assays (ELISA). Disorders which are
CC diagnosed or treated include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides are used to
CC modulate mammalian metabolism, to change mammal's mental state or
CC physical state by influencing biorhythms circadian rhythms, depression
CC tendency for violence tolerance for pain, reproductive capabilities,
CC hormonal or endocrine levels, appetite, libido, memory, stress or other
CC cognitive qualities, as a food additive or preservative, such as to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrates, vitamins, minerals, cofactors or other nutritional
CC components.
XX
SQ Sequence 63 AA;

Query Match 13.1%; Score 58; DB 21; Length 63;
Best Local Similarity 30.4%; Pred. No. 23;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

```
Qy       2 RLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLC-CHRVPSPNSTNLKGHHVRLC 60
           ||| ||| |||:|               ||:   :  |  | : :  || : |  :  |
Db       5 RLLGLSSFLCIIL-------------GLAWTAPASESCGPHPLAAEPSTVILG-AIFPC 49

Qy      61 KPCKLEPEP 69
             :   | | |
Db      50 RTGSLSPAP 58
```

RESULT 15
AAB60725
ID AAB60725 standard; protein; 63 AA.
XX
AC AAB60725;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human secreted protein #24.
XX
KW Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection.
XX
OS Homo sapiens.

XX
PN WO200076531-A1.
XX
PD 21-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15137.
XX
PR 11-JUN-1999; 99US-0138625.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis GA;
XX
DR WPI; 2001-071148/08.
XX
PT Nucleic acids encoding 47 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
PS Claim 11; Page 485; 525pp; English.
XX
CC The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
CC infectious diseases.
XX
SQ Sequence 63 AA;

Query Match 13.1%; Score 58; DB 22; Length 63;
Best Local Similarity 30.4%; Pred. No. 23;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

```
Qy       2 RLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLC-CHRVPSPNSTNLKGHHVRLC 60
           ||| ||| |||:|               ||:   :  |  | : :  || : |  :  |
Db       5 RLLGLSSFLCIIL-------------GLAWTAPASESCGPHPLAAEPSTVILG-AIFPC 49

Qy      61 KPCKLEPEP 69
             :   | | |
Db      50 RTGSLSPAP 58
```

Search completed: May 8, 2003, 21:46:30
Job time : 36 secs

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:47:43 ; Search time 1229 Seconds
(without alignments)
1918.086 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 463318

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113321_7344/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_1319_@runat_07052003_113321_7344 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	59.5	13.5	76	6	A07238	A07238 Probe for v
	2	56	12.7	42	6	AX342232	AX342232 Sequence
	3	53	12.0	79	6	I86930	I86930 Sequence 19
c	4	51.5	11.7	69	9	HSU90998	U90998 Homo sapien
	5	49	11.1	66	10	MUSTCAYJ	M38677 Mouse T-cel
	6	49	11.1	70	10	RNTCRAJ39	Y09178 R.norvegicu
	7	48	10.9	76	9	S75385S18	S75420 collagen ty
	8	47	10.6	51	6	AX161822	AX161822 Sequence
	9	46.5	10.5	48	6	AR077574	AR077574 Sequence
	10	46	10.4	71	6	AX025072	AX025072 Sequence
	11	46	10.4	71	6	AX025094	AX025094 Sequence
c	12	46	10.4	72	6	I71481	I71481 Sequence 22
	13	46	10.4	78	9	HSIGKAB01	X54166 Human BL Ly
	14	45.5	10.3	69	6	AR005307	AR005307 Sequence
	15	45	10.2	39	6	AR077518	AR077518 Sequence
c	16	45	10.2	48	6	AX160227	AX160227 Sequence
c	17	45	10.2	76	9	AF118405	AF118405 Homo sapi
	18	44	10.0	42	6	AR014091	AR014091 Sequence
c	19	44	10.0	50	6	I28909	I28909 Sequence 7
c	20	44	10.0	50	6	I28910	I28910 Sequence 8
	21	44	10.0	51	6	AR077572	AR077572 Sequence
	22	44	10.0	51	6	AR077575	AR077575 Sequence
c	23	44	10.0	66	6	A28858	A28858 Oligonucleo
c	24	44	10.0	79	5	DUKAETC04	M55135 Peking duck
c	25	43	9.7	46	6	A06605	A06605 Oligonucleo
c	26	43	9.7	46	6	A10221	A10221 adapter. 1/
	27	43	9.7	51	6	AX159330	AX159330 Sequence
	28	43	9.7	66	6	AX040385	AX040385 Sequence
	29	43	9.7	72	9	HSADD2S31	AF000111 Homo sapi
c	30	43	9.7	76	1	MCTRFM	X16759 Mycoplasma
c	31	43	9.7	76	1	MYCTRMF	K00312 Mycoplasma
	32	43	9.7	76	6	A07239	A07239 Probe for t
	33	43	9.7	78	6	AR000573	AR000573 Sequence
c	34	43	9.7	78	6	AR000574	AR000574 Sequence
	35	43	9.7	78	6	AR208886	AR208886 Sequence
c	36	43	9.7	78	6	AR208887	AR208887 Sequence
	37	43	9.7	78	6	BD003625	BD003625 A polynuc
c	38	43	9.7	78	6	BD003626	BD003626 A polynuc
c	39	43	9.7	79	6	AR050027	AR050027 Sequence
c	40	42.5	9.6	69	9	HSU91209	U91209 Homo sapien
c	41	42	9.5	30	6	AR022280	AR022280 Sequence
c	42	42	9.5	47	6	AX003724	AX003724 Sequence
	43	42	9.5	51	6	AX159329	AX159329 Sequence
c	44	42	9.5	51	6	AX165793	AX165793 Sequence
	45	42	9.5	52	9	HS19	X68693 H.sapiens m

ALIGNMENTS

RESULT 1

```
A07238
LOCUS       A07238                    76 bp    DNA     linear   PAT 24-AUG-1993
DEFINITION  Probe for vPA-beta.
ACCESSION   A07238
VERSION     A07238.1  GI:413006
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Baldus,B., Donner,P., Schleuning,W.D., Alagon,A., Boidol,W.,
            Kraetzschmar,J.R., Haendler,B.J. and Langer,G.
  TITLE     Novel thrombolytic
  JOURNAL   Patent: EP 0383417-A 1 22-AUG-1990;
            SCHERING AKTIENGESELLSCHAFT
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      14 a     17 c     24 g     21 t
ORIGIN

Alignment Scores:
Pred. No.:              566          Length:         76
Score:                  59.50        Matches:        14
Percent Similarity:     66.67%       Conservative:   4
Best Local Similarity:  51.85%       Mismatches:     7
Query Match:            13.46%       Indels:         2
DB:                     6            Gaps:           1

US-09-599-087B-5 (1-81) x A07238 (1-76)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePheSer-ThrGluGlyLysArgArgPr 28
          ||||||||||:::|||||||||||||   ::::::||||||       |||   |||:::||
Db      1 CTGTTGTGTGTACTGCTGCTTTGTGGAGCAGTCTTCTCGTTGCCCAGGC---AGGAAACC 57

Qy     28 oAlaLysAlaTrpSerGly 34
          |       |||   |||
Db     58 TACAGGCAATTGGCAAGGG 76

RESULT 2
AX342232
LOCUS       AX342232                  42 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 18 from Patent WO0198497.
ACCESSION   AX342232
VERSION     AX342232.1  GI:18151770
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 18 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR primer corresponding to human SECS-1"
BASE COUNT      15 a     10 c      9 g      8 t
ORIGIN

Alignment Scores:
Pred. No.:              642          Length:         42
Score:                  56.00        Matches:        10
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            12.67%       Indels:         0
DB:                     6            Gaps:           0

US-09-599-087B-5 (1-81) x AX342232 (1-42)

Qy     25 LysArgArgProAlaLysAlaTrpSerGly 34
          ||||||||||||||||||||||||||||||
Db     13 AAACGTCGTCCAGCTAAAGCCTGGTCAGGC 42

RESULT 3
I86930
LOCUS       I86930                    79 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 19 from patent US 5702931.
ACCESSION   I86930
VERSION     I86930.1  GI:3206648
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Andrews,W.H., Morser,M.J. and Vilander,L.R.
  TITLE     Mutagenesis methods and compositions
  JOURNAL   Patent: US 5702931-A 19 30-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="unknown"
BASE COUNT      10 a     28 c     31 g     10 t
ORIGIN

Alignment Scores:
Pred. No.:              2.45e+03     Length:         79
Score:                  53.00        Matches:        9
Percent Similarity:     55.00%       Conservative:   2
Best Local Similarity:  45.00%       Mismatches:     9
Query Match:            11.99%       Indels:         0
DB:                     6            Gaps:           0

US-09-599-087B-5 (1-81) x I86930 (1-79)

Qy     29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
          |||      |||:::|||   |||         |||   ||||||   |||   |||:::
Db     15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCTTGCCCGAGT 74

RESULT 4
HSU90998/c
LOCUS       HSU90998                  69 bp    mRNA    linear   PRI 15-JUL-1997
DEFINITION  Homo sapiens clone FE1389 T-cell receptor delta chain (TCRDV1J2)
            mRNA, partial cds.
ACCESSION   U90998
VERSION     U90998.1  GI:2239403
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
            Kagnoff,M.F.
  TITLE     The TCR-delta repertoire in human intestine undergoes
            characteristic changes during fetal to adult development
  JOURNAL   J. Immunol. 158 (12), 5632-5641 (1997)
  MEDLINE   97334214
   PUBMED   9190911
REFERENCE   2  (bases 1 to 69)
  AUTHORS   Holtmeier,W., Witthoeft,T., Hennemann,A., Harland,S.W. and
            Kagnoff,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1997) Department of Medicine, University of
            Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="FE1389"
                     /tissue_type="small intestine"
                     /dev_stage="fetus; 20 weeks gestation"
```

```
     gene            <1. .>69
                     /gene="TCRDV1J2"
     CDS             <1. .>69
                     /gene="TCRDV1J2"
                     /note="rearranged; contains CDR3 domain, 11 amino acids
                     (calculation according to: Rock,E., J.Exp.Med.,
                     179:323-328,1994)"
                     /codon_start=1
                     /product="T-cell receptor delta chain"
                     /protein_id="AAC51561.1"
                     /db_xref="GI:2239404"
                     /translation="LEDSAKYFCALGELGDTVTAQLF"
     V_region        <1. .>69
                     /gene="TCRDV1J2"
BASE COUNT       18 a     14 c     19 g     18 t
ORIGIN

Alignment Scores:
Pred. No.:              2.95e+03        Length:         69
Score:                  51.50           Matches:        9
Percent Similarity:     73.33%          Conservative:   2
Best Local Similarity:  60.00%          Mismatches:     3
Query Match:            11.65%          Indels:         1
DB:                     9               Gaps:           1

US-09-599-087B-5 (1-81) x HSU90998 (1-69)

Qy      39 LeuCysCysHisArgValProSer---ProAsnSerThrAsnLeu 52
           ||||||||||||||:::|||      |||   ||||||   :::
Db      63 TTGTGCTGTCACCGTATCCCCCAGTTCCCCAAGAGCACAAAAGTA 19

RESULT 5
MUSTCAYJ
LOCUS       MUSTCAYJ                   66 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T-cell receptor alpha-chain mRNA J-C region, partial cds.
ACCESSION   M38677
VERSION     M38677.1  GI:201214
KEYWORDS    C-region; J-region; T-cell receptor.
SOURCE      Mouse, cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Yague,J., Blackman,M., Born,W., Marrack,P., Kappler,J. and
            Palmer,E.
  TITLE     The structure of V alpha and J alpha segments in the mouse
  JOURNAL   Nucleic Acids Res. 16 (23), 11355-11364 (1988)
  MEDLINE   89083566
   PUBMED   2849763
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .66
                     /partial
                     /codon_start=1
                     /product="T-cell receptor alpha chain"
                     /protein_id="AAA40195.1"
                     /db_xref="GI:201215"
                     /translation="TTASLGKLQFGTGTQVVVTPDI"
BASE COUNT       18 a     16 c     19 g     13 t
ORIGIN

Alignment Scores:
Pred. No.:              4.86e+03        Length:         66
Score:                  49.00           Matches:        8
Percent Similarity:     80.00%          Conservative:   0
Best Local Similarity:  80.00%          Mismatches:     2
Query Match:            11.09%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x MUSTCAYJ (1-66)

Qy      63 CysLysLeuGluProGluProArgLeuTrp 72
           |||   ||||||   |||||||||||||||
Db      23 TGCAGTTTGGAACAGGAACCCAGGTTGTGG 52

RESULT 6
RNTCRAJ39
LOCUS       RNTCRAJ39                  70 bp    mRNA    linear   ROD 06-JAN-1998
DEFINITION  R.norvegicus TCR alpha joining region, clone-library VA8s2F39.
ACCESSION   Y09178
VERSION     Y09178.1  GI:1669729
KEYWORDS    TCR alpha joining region.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Torres-Nagel,N., Deutschlander,A., Herrmann,T., Arden,B. and
            Hunig,T.
  TITLE     Control of TCR V alpha-mediated positive repertoire selection and
            alloreactivity by differential J alpha usage and CDR3 alpha
            composition
  JOURNAL   Int. Immunol. 9 (10), 1441-1452 (1997)
  MEDLINE   98013626
   PUBMED   9352349
REFERENCE   2  (bases 1 to 70)
  AUTHORS   Torres-Nagel,N.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-1996) N.E. Torres-Nagel, Institut of Virology and
            Immunbiology, Department of Immunology, Versbacherstr. 7, 97078
            Wuerzburg, FRG
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Rattus norvegicus"
                     /strain="LEW.1F"
                     /db_xref="taxon:10116"
                     /haplotype="RT1f"
                     /cell_type="peripheral T cell"
                     /clone_lib="VA8s2F39"
                     /dev_stage="adult"
                     /rearranged
     gene            1. .70
                     /gene="TCRAJ"
     CDS             <1. .>70
                     /gene="TCRAJ"
                     /codon_start=1
                     /protein_id="CAA70377.1"
                     /db_xref="GI:1785816"
                     /translation="AVEIATGFGKLQFGTGTQVVVTP"
     J_segment       <1. .>70
                     /gene="TCRAJ"
                     /note="CDR3 region and J alpha region"
                     /evidence=experimental
BASE COUNT       16 a     13 c     24 g     17 t
ORIGIN

Alignment Scores:
Pred. No.:              5.18e+03        Length:         70
Score:                  49.00           Matches:        8
Percent Similarity:     80.00%          Conservative:   0
Best Local Similarity:  80.00%          Mismatches:     2
Query Match:            11.09%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x RNTCRAJ39 (1-70)

Qy      63 CysLysLeuGluProGluProArgLeuTrp 72
           |||   ||||||   |||||||||||||||
Db      32 TGCAGTTTGGAACAGGAACCCAGGTTGTGG 61

RESULT 7
S75385S18
LOCUS       S75385S18                  76 bp    DNA     linear   PRI 02-JUN-2000
```

```
DEFINITION  collagen type VI alpha 1(VI) chain (exons 3-19 and exon/intron
            junctions) [human, Genomic, 76 nt, segment 18 of 18].
ACCESSION   S75420
VERSION     S75420.1  GI:241983
KEYWORDS    .
SEGMENT     18 of 18
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Saitta,B., Wang,Y.M., Renkart,L., Zhang,R.Z., Pan,T.C., Timpl,R.
            and Chu,M.L.
  TITLE     The exon organization of the triple-helical coding regions of the
            human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
  JOURNAL   Genomics 11 (1), 145-153 (1991)
  MEDLINE   92112205
   PUBMED   1765372
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 75420] from the original journal article.
            This sequence comes from Figure 2.
            Map location: chromosome 21.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     mRNA            join(S75385.1:<30. .74,S75388.1:31. .57,S75390.1:31. .57,
                     S75392.1:31. .75,S75394.1:31. .84,S75396.1:31. .93,
                     S75398.1:31. .93,S75400.1:31. .84,S75402.1:31. .66,
                     S75404.1:31. .93,S75406.1:31. .93,S75408.1:31. .93,
                     S75410.1:31. .93,S75412.1:31. .81,S75414.1:31. .66,
                     S75416.1:31. .93,S75418.1:31. .96,31. .>66)
                     /product="collagen type VI alpha 1(VI) chain"
     CDS             join(S75385.1:<30. .74,S75388.1:31. .57,S75390.1:31. .57,
                     S75392.1:31. .75,S75394.1:31. .84,S75396.1:31. .93,
                     S75398.1:31. .93,S75400.1:31. .84,S75402.1:31. .66,
                     S75404.1:31. .93,S75406.1:31. .93,S75408.1:31. .93,
                     S75410.1:31. .93,S75412.1:31. .81,S75414.1:31. .66,
                     S75416.1:31. .93,S75418.1:31. .96,31. .>66)
                     /gene="collagen type VI alpha 1(VI) chain"
                     /note="This sequence comes from Figure 2; conceptual
                     translation differs from the translation provided in the
                     manuscript"
                     /codon_start=1
                     /product="collagen type VI alpha 1(VI) chain"
                     /protein_id="AAB20835.2"
                     /db_xref="GI:8176524"
                     /translation="GRPGDLGPVGYQGMKGEKGSRGEKGSRGPKGYKGEKGKRGIDGV
                     DGVKGEMGYPGLPGCKGSPGFDGIQGPPGPKGDPGAFGLKGEKGEPGADGEAGRPGAR
                     GPSGDEGPAGEPGPPGEKGEAGDEGNPGPDGAPGERGGPGERGPRGTPGTRGPRGDPG
                     EAGPQGDQGREGPVGVPGDPGEAGPIGPKGYRGDEGPPGSEGARGAPGPAGPPGDPGL
                     MGERGEDGPAGNGTEGFPGFPGYPGNRGAPGINGTKGYPGLKGDEGEAGDPGDDNNDI
                     APRGVKGAKGYRGPEGPQGPPGHQGPPGPD"
     exon            31. .>66
                     /number=20
BASE COUNT       15 a     26 c     24 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:                7.04e+03        Length:          76
Score:                    48.00           Matches:         9
Percent Similarity:       59.09%          Conservative:    4
Best Local Similarity:    40.91%          Mismatches:      9
Query Match:              10.86%          Indels:          0
DB:                       9               Gaps:            0

US-09-599-087B-5 (1-81) x S75385S18 (1-76)

Qy      15 LeuCysPheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
           |||   :::|||:::     |||        |||   |||      ||||||:::|||
Db       6 CTCCTCTACTCCGTTTCTCGGACAGGGACCCCCAGGACACCAAGGACCGCCTGGGCCGGA 65

Qy      35 ArgArg 36
           ||:::
Db      66 CGTAAG 71

RESULT 8
AX161822
LOCUS       AX161822                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5150 from Patent WO0140521.
ACCESSION   AX161822
VERSION     AX161822.1  GI:14543153
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5150 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5149 is other entry)
                     Accession number cg43984006"
BASE COUNT       14 a     13 c     13 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:                5.7e+03         Length:          51
Score:                    47.00           Matches:         8
Percent Similarity:       72.73%          Conservative:    0
Best Local Similarity:    72.73%          Mismatches:      3
Query Match:              10.63%          Indels:          0
DB:                       6               Gaps:            0

US-09-599-087B-5 (1-81) x AX161822 (1-51)

Qy      39 LeuCysCysHisArgValProSerProAsnSer 49
           |||||||||   |||   ||||||   ||||||
Db      19 TTGTGCTGTGAACGAGCACCGTCAGAGAATTCC 51

RESULT 9
AR077574
LOCUS       AR077574                 48 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 289 from patent US 5962255.
ACCESSION   AR077574
VERSION     AR077574.1  GI:10004320
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Griffiths,A.David., Williams,S.Cameron., Waterhouse,P.Michael.,
            Nissim,A., Winter,G.Paul., Johnson,K.Stuart. and
            Smith,A.John.Hammond.
  TITLE     Methods for producing recombinant vectors
  JOURNAL   Patent: US 5962255-A 289 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT        8 a     18 c     11 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:                5.95e+03        Length:          48
Score:                    46.50           Matches:         9
Percent Similarity:       66.67%          Conservative:    3
Best Local Similarity:    50.00%          Mismatches:      3
Query Match:              10.52%          Indels:          3
```

```
DB:                    6              Gaps:          1

US-09-599-087B-5 (1-81) x AR077574 (1-48)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
           ||||||  ||||||   :::   ||||||         |||:::|||:::|||
Db       3 TGGTCCCTCCGCCGAATACAGTGCTGCTGT---------CCCACGCCTGACAGT 47

RESULT 10
AX025072
LOCUS       AX025072                  71 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 18 from Patent WO0031127.
ACCESSION   AX025072
VERSION     AX025072.1  GI:10184987
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Hope,G. and Mclauchlan,J.
  TITLE     Protein fragments for use in protein targeting
  JOURNAL   Patent: WO 0031127-A 18 02-JUN-2000;
            MEDICAL RES COUNCIL (GB) ; HOPE GRAHAM (GB) ; MCLAUCHLAN JOHN (GB)
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide used in mutagenesis"
BASE COUNT       13 a     20 c     22 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              1.01e+04       Length:        71
Score:                  46.00          Matches:       9
Percent Similarity:     68.75%         Conservative:  2
Best Local Similarity:  56.25%         Mismatches:    5
Query Match:            10.41%         Indels:        0
DB:                     6              Gaps:          0

US-09-599-087B-5 (1-81) x AX025072 (1-71)

Qy      25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           :::||||||||||::         ||||||||||||      ||||||
Db      19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66

RESULT 11
AX025094
LOCUS       AX025094                  71 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 18 from Patent WO0031532.
ACCESSION   AX025094
VERSION     AX025094.1  GI:10185008
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Hope,G. and Mclauchlan,J.
  TITLE     Methods of identifying viral therapeutics
  JOURNAL   Patent: WO 0031532-A 18 02-JUN-2000;
            MEDICAL RES COUNCIL (GB) ; HOPE GRAHAM (GB) ; MCLAUCHLAN JOHN (GB)
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide used in mutagenesis"
BASE COUNT       13 a     20 c     22 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              1.01e+04       Length:        71
Score:                  46.00          Matches:       9
Percent Similarity:     68.75%         Conservative:  2
Best Local Similarity:  56.25%         Mismatches:    5
Query Match:            10.41%         Indels:        0
DB:                     6              Gaps:          0

US-09-599-087B-5 (1-81) x AX025094 (1-71)

Qy      25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           :::||||||||||::         ||||||||||||      ||||||
Db      19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66

RESULT 12
I71481/c
LOCUS       I71481                    72 bp    DNA     linear   PAT 03-APR-1998
DEFINITION  Sequence 22 from patent US 5681942.
ACCESSION   I71481
VERSION     I71481.1  GI:3007616
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Buchwald,M., Strathdee,C.A., Wevrick,R. and Mathew,C.George.Porter.
  TITLE     Fanconi Anemia Type C gene
  JOURNAL   Patent: US 5681942-A 22 28-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="unknown"
BASE COUNT       12 a     16 c     13 g     31 t
ORIGIN

Alignment Scores:
Pred. No.:              1.03e+04       Length:        72
Score:                  46.00          Matches:       7
Percent Similarity:     76.92%         Conservative:  3
Best Local Similarity:  53.85%         Mismatches:    3
Query Match:            10.41%         Indels:        0
DB:                     6              Gaps:          0

US-09-599-087B-5 (1-81) x I71481 (1-72)

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
           |||||||||:::   :::|||   |||   |||:::|||
Db      71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 13
HSIGKAB01
LOCUS       HSIGKAB01                 78 bp    mRNA    linear   PRI 04-AUG-1992
DEFINITION  Human BL Ly66 mRNA for abnormal Ig kappa light chain (1).
ACCESSION   X54166
VERSION     X54166.1  GI:33218
KEYWORDS    Ig kappa light chain; Ig light chain; immunoglobulin.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   COGNE,M.C.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1990) Cogne M.C.C., CNRS URA 1172, Lab.
            Immunologie Moleculaire, Faculte des Sciences, F-86022 Poitiers,
            france
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Cogne,M., Mounir,S., Mahdi,T., Preud'homme,J.L., Nau,F. and
            Guglielmi,P.
  TITLE     Production of an abnormal mu chain with a shortened VHIV subgroup
            variable region in a Burkitt's lymphoma cell line
  JOURNAL   Mol. Immunol. 27 (9), 929-934 (1990)
  MEDLINE   91015095
   PUBMED   2120579
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="Ly66"
```

```
                     /cell_type="B-cell"
                     /tissue_type="Burkitt lymphoma"
     mRNA            1. .78
                     /note="abnormal Ig L chain 0.85kb transcript"
                     /evidence=experimental
     exon            14. .68
                     /note="V kappa 1 leader peptide exon"
                     /number=1
     exon            69. .>78
                     /note="C kappa exon, corresponds to Km3  allele sequence"
                     /number=2
BASE COUNT       13 a     24 c     23 g     18 t
ORIGIN

Alignment Scores:
Pred. No.:              1.12e+04        Length:         78
Score:                  46.00           Matches:        10
Percent Similarity:     76.47%          Conservative:   3
Best Local Similarity:  58.82%          Mismatches:     4
Query Match:            10.41%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x HSIGKAB01 (1-78)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPhe 17
           |||   :::   ||||||:::   ||||||   :::|||||||||||||||
Db      14 ATGGACATGAGAGTCCTCGCTCAGCTCCTGGGGCTCCTGCTGCTCTGTTTC 64

RESULT 14
AR005307
LOCUS       AR005307                  69 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 37 from patent US 5747660.
ACCESSION   AR005307
VERSION     AR005307.1  GI:3966186
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Orlicky,D.J.
  TITLE     Nucleic acid encoding prostaglandin F.sub.2.alpha. receptor
            regulatory protein
  JOURNAL   Patent: US 5747660-A 37 05-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="unknown"
BASE COUNT        6 a     25 c     31 g      7 t
ORIGIN

Alignment Scores:
Pred. No.:              1.1e+04         Length:         69
Score:                  45.50           Matches:        10
Percent Similarity:     54.55%          Conservative:   2
Best Local Similarity:  45.45%          Mismatches:     7
Query Match:            10.29%          Indels:         3
DB:                     6               Gaps:           1

US-09-599-087B-5 (1-81) x AR005307 (1-69)

Qy      25 LysArgArgProAla---------LysAlaTrpSerGlyArgArgThrArgLeuCysCys 41
           :::||||||   |||         :::||||||      ||||||         ||||||
Db       3 AGGCGGAGAGTCGCTCCCGCCGGCCGAGCATGGGGCGCCCGGCGCCGAGGCCGCTGCTGC 62

Qy      42 HisArg 43
             |||
Db      63 TGGCGC 68

RESULT 15
AR077518
LOCUS       AR077518                  39 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 233 from patent US 5962255.
ACCESSION   AR077518
VERSION     AR077518.1  GI:10004264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Griffiths,A.David., Williams,S.Cameron., Waterhouse,P.Michael.,
            Nissim,A., Winter,G.Paul., Johnson,K.Stuart. and
            Smith,A.John.Hammond.
  TITLE     Methods for producing recombinant vectors
  JOURNAL   Patent: US 5962255-A 233 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT        7 a     13 c     10 g      9 t
ORIGIN

Alignment Scores:
Pred. No.:              6.61e+03        Length:         39
Score:                  45.00           Matches:        8
Percent Similarity:     66.67%          Conservative:   0
Best Local Similarity:  66.67%          Mismatches:     4
Query Match:            10.18%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x AR077518 (1-39)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArg 43
           ||||||   ||||||         ||||||||||
Db       3 TGGTCCCTCCGCCGAATACCAGGCTGCTGTCATAGG 38


Search completed: May  8, 2003, 22:13:41
Job time : 1232 secs
```

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 22:13:48 ; Search time 1348 Seconds
(without alignments)
973.171 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 227956

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113852_7677/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_1525_@runat_07052003_113852_7677 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	50.5	11.4	60	17	AZ918344	AZ918344 1006004B0
c 2	50.5	11.4	69	17	BH639950	BH639950 1008032H0
c 3	49	11.1	72	14	BQ394194	BQ394194 NISC_ng07
4	49	11.1	76	17	AZ918366	AZ918366 1006004B0
c 5	49	11.1	79	9	AI877052	AI877052 uc56b11.r
6	48	10.9	80	10	AV676540	AV676540 AV676540
c 7	47	10.6	67	17	AZ605110	AZ605110 1M0426I15
8	47	10.6	73	10	AW600148	AW600148 SWL4CAK10
c 9	47	10.6	74	9	AA739046	AA739046 vv66d04.r
c 10	47	10.6	74	13	BM283913	BM283913 ki35b04.y
c 11	46.5	10.5	79	9	AI468972	AI468972 ti43h02.x
12	46	10.4	50	9	AU105245	AU105245 AU105245
c 13	46	10.4	70	14	BQ092063	BQ092063 fz36f08.x
c 14	46	10.4	70	14	BQ092271	BQ092271 fz39e09.x
c 15	46	10.4	73	14	H22549	H22549 yn69h12.r1
16	46	10.4	77	9	AA049559	AA049559 mj35b12.r
c 17	45.5	10.3	77	14	BQ666094	BQ666094 pb35g10.y
c 18	45.5	10.3	77	14	BQ811778	BQ811778 1030025E1
19	45	10.2	67	13	BM518515	BM518515 ki93f08.y
c 20	45	10.2	70	17	AZ992067	AZ992067 2M0276O01
21	45	10.2	72	14	BQ757145	BQ757145 EBem10_SQ
c 22	45	10.2	72	17	AZ493456	AZ493456 1M0328G07
c 23	45	10.2	72	17	AZ921678	AZ921678 1006031C0
24	45	10.2	73	9	AA790708	AA790708 vw18f08.r
25	45	10.2	79	9	AA498761	AA498761 vh32d12.r
26	45	10.2	80	9	AU257386	AU257386 AU257386
27	44	10.0	74	13	BM097862	BM097862 EBem04_SQ
28	44	10.0	76	9	AU244358	AU244358 AU244358
29	44	10.0	78	6	BQ241316	Bq241316 TaE05006F
30	43.5	9.8	66	17	BH620911	BH620911 1007099G0
c 31	43.5	9.8	67	17	AZ808082	AZ808082 2M0071I24
c 32	43.5	9.8	72	14	R33083	R33083 yh77c06.s1
33	43	9.7	40	9	AI589916	AI589916 tm81h06.x
34	43	9.7	48	17	TA176E12P	AL475046 T. brucei
35	43	9.7	51	10	AW232755	AW232755 fj30a01.x
c 36	43	9.7	58	9	AA948249	AA948249 oq33g08.s
37	43	9.7	58	17	AZ835896	AZ835896 2M0130C22
c 38	43	9.7	62	17	FR0012583	AL003835 F.rubripe
39	43	9.7	64	10	AW626440	AW626440 SWOvAFCAP
40	43	9.7	64	10	AW626725	AW626725 SWOvAFCAP
41	43	9.7	64	12	BF118532	BF118532 SWOvL3CAN
42	43	9.7	70	9	AU258504	AU258504 AU258504
c 43	43	9.7	72	10	AV911900	AV911900 AV911900
c 44	43	9.7	75	9	AL643544	AL643544 AL643544
c 45	43	9.7	75	13	BI753134	BI753134 603025986

ALIGNMENTS

RESULT 1
AZ918344/c
LOCUS AZ918344 60 bp DNA linear GSS 17-DEC-2001
DEFINITION 1006004B02.x1 1006 - RescueMu Grid G Zea mays genomic, DNA sequence.
ACCESSION AZ918344
VERSION AZ918344.1 GI:13387628
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 60)
AUTHORS Walbot,V.

```
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006004  row: 36
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        5 a     17 c     26 g     12 t
ORIGIN

Alignment Scores:
Pred. No.:                9.11e+03       Length:        60
Score:                    50.50          Matches:       10
Percent Similarity:       78.57%         Conservative:  1
Best Local Similarity:    71.43%         Mismatches:    2
Query Match:              11.43%         Indels:        1
DB:                       17             Gaps:          1

US-09-599-087B-5 (1-81) x AZ918344 (1-60)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerPro 47
           |||   ||||||||||||   |||||||||   |||:::|||
Db      43 GGCCACCGCACACGGCTCATCTGTCACCGG---CCCACGCCA 5

RESULT 2
BH639950/c
LOCUS       BH639950                 69 bp    DNA     linear   GSS 14-FEB-2002
DEFINITION  1008032H07.2EL_y1 1008 - RescueMu Grid I Zea mays genomic, DNA
            sequence.
ACCESSION   BH639950
VERSION     BH639950.1  GI:18665756
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008032  row: 22
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1008 - RescueMu Grid I"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        7 a     30 c     26 g      6 t
ORIGIN

Alignment Scores:
Pred. No.:                1.11e+04       Length:        69
Score:                    50.50          Matches:       10
Percent Similarity:       66.67%         Conservative:  2
Best Local Similarity:    55.56%         Mismatches:    5
Query Match:              11.43%         Indels:        1
DB:                       17             Gaps:          1

US-09-599-087B-5 (1-81) x BH639950 (1-69)

Qy      25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHis 42
           :::|||||||||||||:::   |||      |||   ||||||   |||   |||
Db      62 CGCCGGAGGCCGGCGCGGCGGTGGCGC---CGGCTGACCCGGAGTTGCCTCCAC 12

RESULT 3
BQ394194/c
LOCUS       BQ394194                 72 bp    mRNA    linear   EST 22-MAY-2002
DEFINITION  NISC_ng07g11.x1 NICHD XGC Emb6 Silurana tropicalis cDNA clone
            IMAGE:5382740 3', mRNA sequence.
ACCESSION   BQ394194
VERSION     BQ394194.1  GI:21081881
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   NIH-XCG http://image.llnl.gov/image/html/xenopuslib_info.shtml.
  TITLE     National Institute of Child Health and Human Development, National
            Cancer Insitute, Xenopus Gene Collection
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11974  row: N  column: 21
            Seq primer: -21M13 forward primer (ABI).
```

```
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:5382740"
                     /clone_lib="NICHD XGC Emb6"
                     /tissue_type="neurula"
                     /dev_stage="embryo, stages 14-19"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6.ccdb; Site_1: NotI; Site_2:
                     EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average
                     insert size 2.1 kb. Constructed by Invitrogen. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT       13 a     14 c      5 g     40 t
ORIGIN

Alignment Scores:
Pred. No.:                1.7e+04        Length:         72
Score:                    49.00          Matches:        9
Percent Similarity:       65.22%         Conservative:   6
Best Local Similarity:    39.13%         Mismatches:     8
Query Match:              11.09%         Indels:         0
DB:                       14             Gaps:           0

US-09-599-087B-5 (1-81) x BQ394194 (1-72)

Qy      17 PheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArg 36
           ||||||:::|||      :::|||||||:::::      |||         |||   |||
Db      69 TTTAGTGTTTTTGAAAAAAAAGGGAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGC 10

Qy      37 ThrArgLeu 39
           :::|||:::
Db       9 TCTAGAGTA 1

RESULT 4
AZ918366
LOCUS       AZ918366                 76 bp    DNA     linear   GSS 17-DEC-2001
DEFINITION  1006004B05.2EL_x3 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ918366
VERSION     AZ918366.1  GI:13387650
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006004  row: 36
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       15 a     24 c     22 g     15 t
ORIGIN

Alignment Scores:
Pred. No.:                1.84e+04       Length:         76
Score:                    49.00          Matches:        10
Percent Similarity:       57.89%         Conservative:   1
Best Local Similarity:    52.63%         Mismatches:     2
Query Match:              11.09%         Indels:         6
DB:                       17             Gaps:           1

US-09-599-087B-5 (1-81) x AZ918366 (1-76)

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           |||   :::||||||||||                    ||||||||||||||||||
Db      38 CACTCAATGCCATCGCCA----------------CCCCACCACGTCAGATTATGT 76

RESULT 5
AI877052/c
LOCUS       AI877052                 79 bp    mRNA    linear   EST 21-JUL-1999
DEFINITION  uc56b11.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:1429629 5' similar to TR:O35500 O35500 KINESIN LIGHT CHAIN
            ISOFORM D. ;, mRNA sequence.
ACCESSION   AI877052
VERSION     AI877052.1  GI:5551101
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:913697
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1429629"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
```

rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 15 a 23 c 25 g 16 t
ORIGIN

Alignment Scores:
Pred. No.: 1.95e+04 Length: 79
Score: 49.00 Matches: 10
Percent Similarity: 47.83% Conservative: 1
Best Local Similarity: 43.48% Mismatches: 12
Query Match: 11.09% Indels: 0
DB: 9 Gaps: 0

US-09-599-087B-5 (1-81) x AI877052 (1-79)

```
Qy      40 CysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeu 59
           |||         |||||||||   :::|||   |||      |||          |||
Db      78 TGCGCAGGAGCAGTTCCTTCTCTAAGCTCCGGGAATCCATCAGGCGAGGAAGCGAGAAGC 19

Qy      60 CysLysPro 62
           |||   |||
Db      18 TGTTCTCCC 10
```

RESULT 6
AV676540
LOCUS AV676540 80 bp mRNA linear EST 05-OCT-2000
DEFINITION AV676540 Nori Satoh unpublished cDNA library Ciona intestinalis cDNA clone rcitble23 3', mRNA sequence.
ACCESSION AV676540
VERSION AV676540.1 GI:10114539
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 80)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES Location/Qualifiers
source 1. .80
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="rcitble23"
/clone_lib="Nori Satoh unpublished cDNA library"
/tissue_type="whole animal"
/dev_stage="tailbud"
/note="Vector: pBluescript SK"
BASE COUNT 20 a 19 c 15 g 26 t
ORIGIN

Alignment Scores:
Pred. No.: 2.53e+04 Length: 80
Score: 48.00 Matches: 8
Percent Similarity: 68.42% Conservative: 5
Best Local Similarity: 42.11% Mismatches: 4
Query Match: 10.86% Indels: 2
DB: 10 Gaps: 1

US-09-599-087B-5 (1-81) x AV676540 (1-80)

```
Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArg------ValProSerProAsn 48
           |||   |||:::|||      ||||||   ::::::      :::||||||||||:::
Db      15 TGGGGAGGTAAAAGATGCTCCTTGTGTTCGTACCAATATGTTATTCCAAGCCCAGAT 71
```

RESULT 7
AZ605110/c
LOCUS AZ605110 67 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0426I15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0426I15 F, DNA sequence.
ACCESSION AZ605110
VERSION AZ605110.1 GI:11727300
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 67)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0426 row: I column: 15
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 67.
FEATURES Location/Qualifiers
source 1. .67
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0426I15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 15 a 17 c 22 g 13 t
ORIGIN

Alignment Scores:
Pred. No.: 2.49e+04 Length: 67
Score: 47.00 Matches: 8
Percent Similarity: 70.59% Conservative: 4
Best Local Similarity: 47.06% Mismatches: 5
Query Match: 10.63% Indels: 0
DB: 17 Gaps: 0

US-09-599-087B-5 (1-81) x AZ605110 (1-67)

```
Qy      44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||  ||||||  :::::::::  |||  |||  :::|||||||
Db      66 GTGCCACCCCCCAACCCCTCTTCTGTGCTAGGGGTCCATTCAAAGCTTTGT 16

RESULT 8
AW600148
LOCUS       AW600148                 73 bp    mRNA    linear   EST 22-MAR-2000
DEFINITION  SWL4CAK10E07SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SWL4CAK10E07 5', mRNA sequence.
ACCESSION   AW600148
VERSION     AW600148.1  GI:7287661
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SWL4CAK10E07"
                     /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                     /dev_stage="larval stage four"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from L4s isolated from the peritoneal
                     cavity of jirds and converted to double-stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNA pol I. The library has 2.7 x 105
                     independent recombinants and the average insert size is
                     approx. 1050bp. The library was constructed by Michelle
                     Lizotte-Waniewski. The library is available from Dr. S.A.
                     Williams, email:genome@neal.smith.edu."
BASE COUNT       21 a     11 c     10 g     31 t
ORIGIN

Alignment Scores:
Pred. No.:              2.82e+04        Length:         73
Score:                  47.00           Matches:        10
Percent Similarity:     70.59%          Conservative:   2
Best Local Similarity:  58.82%          Mismatches:     5
Query Match:            10.63%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x AW600148 (1-73)

Qy       6 LeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThr 22
           ||||||   :::   ||||||   ||||||   ||||||:::|||   |||
Db       3 CTATCACTAGTGTTTTGCATTAGTTTGCTATCTTTTTCAATGTTTGAAACA 53

RESULT 9
AA739046/c
LOCUS       AA739046                 74 bp    mRNA    linear   EST 14-JAN-1998
DEFINITION  vv66d04.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1227367 5' similar to gb:M73436 Mouse ribosomal protein S4
            (MOUSE);, mRNA sequence.
ACCESSION   AA739046
VERSION     AA739046.1  GI:2775852
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:652959
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1227367"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     24 c     15 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              2.88e+04        Length:         74
Score:                  47.00           Matches:        8
Percent Similarity:     83.33%          Conservative:   2
Best Local Similarity:  66.67%          Mismatches:     2
Query Match:            10.63%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x AA739046 (1-74)

Qy      47 ProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||||||||||::::::   ||||||   ||||||
Db      54 CCGAATAGTACGAGCATCATGGGTCACCAGGTGCGG 19

RESULT 10
BM283913/c
LOCUS       BM283913                 74 bp    mRNA    linear   EST 28-DEC-2001
DEFINITION  ki35b04.y1 Ascaris suum female gonad GZ SL1 TOPO v1 Murphy
            Chiapelli McCarter Ascaris suum cDNA 5', mRNA sequence.
ACCESSION   BM283913
VERSION     BM283913.1  GI:17992955
KEYWORDS    EST.
SOURCE      pig roundworm.
  ORGANISM  Ascaris suum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
            ; Ascarididae; Ascaris.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
```

```
          , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
          ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
          , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
          Wilson,R.
  TITLE   The Washington Univ. Nematode EST Project, 1999
  JOURNAL Unpublished (1999)
COMMENT   Contact: McCarter JP
          The Washington Univ. Nematode EST Project, 1999
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          The library was constructed by Claire Murphy, Brandi Chiapelli, and
          Dr. James McCarter at Washington University, St. Louis. DNA
          Sequencing by: Washington University Genome Sequencing Center
          Seq primer: -40RP from Gibco.
FEATURES            Location/Qualifiers
     source         1. .74
                    /organism="Ascaris suum"
                    /db_xref="taxon:6253"
                    /clone_lib="Ascaris suum female gonad GZ SL1 TOPO v1
                    Murphy Chiapelli McCarter"
                    /sex="Female"
                    /tissue_type="Dissected female gonad (GZ-germinal zone,
                    most distal region of gonad)"
                    /dev_stage="Adult"
                    /lab_host="DH10B"
                    /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                    Site_2: EcoRI; The library was constructed by Claire
                    Murphy, Brandi Chiapelli,  and Dr. James McCarter at
                    Washington University, St. Louis. Oligo(dT)-SL1  PCR based
                    library. Ascaris suum female gonad GZ cDNA PCR products of
                    size  >400 nucleotides containing SL1 on the 5' end and
                    oligo(dT) on the 3' end  were non-directionally cloned
                    into pCRII-TOPO(Invitrogen) following the  Topo TA cloning
                    protocol. Dissected nematode tissues were  provided by Dr.
                    Alan Scott (ascott@jhsph.edu) of the School of Public
                    Hygiene  and Public Health at John Hopkins University in
                    Baltimore, MD."
BASE COUNT      30 a      9 c     22 g     13 t
ORIGIN

Alignment Scores:
Pred. No.:                2.88e+04        Length:         74
Score:                    47.00           Matches:        10
Percent Similarity:       73.33%          Conservative:   1
Best Local Similarity:    66.67%          Mismatches:     4
Query Match:              10.63%          Indels:         0
DB:                       13              Gaps:           0

US-09-599-087B-5 (1-81) x BM283913 (1-74)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThrGlu 23
          |||   |||   ||||||||||||||||   |||   |||:::|||
Db     61 CTCCACTGCTTTCTCCTCCTTTGCTTCTTCATCACATCTTCGGAA 17

RESULT 11
AI468972/c
LOCUS       AI468972                 79 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  ti43h02.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2133267 3'
            similar to WP:F59E12.9 CE11534 ;, mRNA sequence.
ACCESSION   AI468972
VERSION     AI468972.1  GI:4331062
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11547-015
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES            Location/Qualifiers
     source         1. .79
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2133267"
                    /clone_lib="NCI_CGAP_Lym12"
                    /tissue_type="lymphoma, follicular mixed small and large
                    cell"
                    /lab_host="DH10B"
                    /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                    SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                    Oligo dT.  Average insert size 1.25 kb. Life Technologies
                    catalog #: 11547-015"
BASE COUNT      12 a     35 c     31 g      1 t
ORIGIN

Alignment Scores:
Pred. No.:                3.57e+04        Length:         79
Score:                    46.50           Matches:        10
Percent Similarity:       54.55%          Conservative:   2
Best Local Similarity:    45.45%          Mismatches:     7
Query Match:              10.52%          Indels:         3
DB:                       9               Gaps:           1

US-09-599-087B-5 (1-81) x AI468972 (1-79)

Qy     58 ArgLeuCysLysProCysLys---------LeuGluProGluProArgLeuTrpValVal 74
          |||   |||   |||                 :::   |||   |||||||:::|||||||
Db     77 CGTTGGTGTGGGCCCCCCCCCCAGGGGGGGGGTGGCCCCCCCCCCGCGTGTTTGGGTGCCC 18

Qy     75 ProGly 76
          ||||||
Db     17 CCCGGG 12

RESULT 12
AU105245
LOCUS       AU105245                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105245 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            COL05671, mRNA sequence.
ACCESSION   AU105245
VERSION     AU105245.1  GI:13554766
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
```

```
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL05671"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        2 a     19 c     18 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:              2.08e+04        Length:         50
Score:                  46.00           Matches:        9
Percent Similarity:     73.33%          Conservative:   2
Best Local Similarity:  60.00%          Mismatches:     4
Query Match:            10.41%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x AU105245 (1-50)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSer 46
           :::|||   ||||||  ||||||  |||   |||:::||||||
Db       5 TTTTCCGCCCGGCGCTGTCGGCTGCTGTGCCCGCGGCTCCCCAGT 49

RESULT 13
BQ092063/c
LOCUS       BQ092063                 70 bp    mRNA    linear   EST 08-APR-2002
DEFINITION  fz36f08.x1 zebrafish fin day1 regeneration Danio rerio cDNA clone
            5905287 3', mRNA sequence.
ACCESSION   BQ092063
VERSION     BQ092063.1  GI:20072973
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: T7 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="5905287"
                     /clone_lib="zebrafish fin day1 regeneration"
                     /sex="mixed male and female"
                     /tissue_type="1 day fin regenerates"
                     /lab_host="E. coli XLOLR"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; 1st
                     strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed
                     by second strand synthesis, and ligated to 5' adapter (5'
                     )-aattcggcacgag- 3', 3'-gccgtgctc-5'.  cDNA was cloned
                     directionally (EcoRI/XhoI) into Stratagene Zap-express
                     lambda phage arms.  Mass invivo excision done to obtain
                     inserts in pBK-CMV phagemid."
BASE COUNT        5 a     19 c     11 g     35 t
ORIGIN

Alignment Scores:
Pred. No.:              3.38e+04        Length:         70
Score:                  46.00           Matches:        8
Percent Similarity:     66.67%          Conservative:   2
Best Local Similarity:  53.33%          Mismatches:     5
Query Match:            10.41%          Indels:         0
DB:                     14              Gaps:           0

US-09-599-087B-5 (1-81) x BQ092063 (1-70)

Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
           |||:::   ||||||         ||||||   :::||||||||||
Db      63 GGAAGAGAGAGGCCACAAACACAGTGGTCGTTCAAGAGGACAAGA 19

RESULT 14
BQ092271/c
LOCUS       BQ092271                 70 bp    mRNA    linear   EST 08-APR-2002
DEFINITION  fz39e09.x1 zebrafish fin day1 regeneration Danio rerio cDNA clone
            5905624 3', mRNA sequence.
ACCESSION   BQ092271
VERSION     BQ092271.1  GI:20073488
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: T7 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="5905624"
                     /clone_lib="zebrafish fin day1 regeneration"
                     /sex="mixed male and female"
                     /tissue_type="1 day fin regenerates"
                     /lab_host="E. coli XLOLR"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; 1st
                     strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed
                     by second strand synthesis, and ligated to 5' adapter (5'
                     )-aattcggcacgag- 3', 3'-gccgtgctc-5'.  cDNA was cloned
```

directionally (EcoRI/XhoI) into Stratagene Zap express lambda phage arms. Mass invivo excision done to obtain inserts in pBK-CMV phagemid."
BASE COUNT 5 a 19 c 11 g 35 t
ORIGIN

Alignment Scores:

Pred. No.:	3.38e+04	Length:	70
Score:	46.00	Matches:	8
Percent Similarity:	66.67%	Conservative:	2
Best Local Similarity:	53.33%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	14	Gaps:	0

US-09-599-087B-5 (1-81) x BQ092271 (1-70)

```
Qy     24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
          |||:::   ||||||         ||||||   :::||||||||||
Db     63 GGAAGAGAGAGGCCACAAACACAGTGGTCGTTCAAGAGGACAAGA 19
```

RESULT 15
H22549/c
LOCUS H22549 73 bp mRNA linear EST 06-JUL-1995
DEFINITION yn69h12.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone IMAGE:173735 5' similar to gb:X56009 GUANINE NUCLEOTIDE-BINDING PROTEIN G(S), ALPHA SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION H22549
VERSION H22549.1 GI:891244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 73)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 755
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 755 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .73
/organism="Homo sapiens"
/db_xref="GDB:3835770"
/db_xref="taxon:9606"
/clone="IMAGE:173735"
/clone_lib="Soares adult brain N2b5HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: brain; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."
BASE COUNT 20 a 21 c 17 g 10 t 5 others
ORIGIN

Alignment Scores:

Pred. No.:	3.59e+04	Length:	73
Score:	46.00	Matches:	9
Percent Similarity:	47.37%	Conservative:	0
Best Local Similarity:	47.37%	Mismatches:	10
Query Match:	10.41%	Indels:	0
DB:	14	Gaps:	0

US-09-599-087B-5 (1-81) x H22549 (1-73)

```
Qy     16 CysPheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
          |||  |||      |||      |||        ||||||     |||||||||
Db     59 TGTTGNTCCAGATGCTCTTNAAGAGGTTCAGAGCCTCCTGCAGACGGTTGGTCTGGT 3
```

Search completed: May 8, 2003, 23:05:56
Job time : 1349 secs

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:43:48 ; Search time 30 Seconds
(without alignments)
556.327 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 69166

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56.5	12.8	57	4	Q96HV2	Q96hv2 homo sapien
2	56.5	12.8	78	10	Q8VZQ7	Q8vzq7 arabidopsis
3	56	12.7	69	5	Q9BKX4	Q9bkx4 caenorhabdi
4	54.5	12.3	75	16	Q9K1L9	Q9k1l9 neisseria m
5	53	12.0	64	16	Q8UE52	Q8ue52 agrobacteri
6	53	12.0	78	12	Q8VAU9	Q8vau9 white spot
7	52.5	11.9	69	12	Q8QN69	Q8qn69 ectocarpus
8	51	11.5	51	16	Q985E2	Q985e2 rhizobium l
9	51	11.5	79	16	Q9KKN6	Q9kkn6 vibrio chol
10	50	11.3	42	8	Q32989	Q32989 pinus thunb
11	50	11.3	59	12	Q9Q6Q2	Q9q6q2 grapevine l
12	50	11.3	76	12	Q91KH1	Q91kh1 hepatitis c
13	49.5	11.2	80	11	Q9JHY4	Q9jhy4 mus musculu
14	48.5	11.0	71	5	Q9VU66	Q9vu66 drosophila
15	48	10.9	61	11	Q61910	Q61910 mus musculu
16	48	10.9	69	2	Q9ZC27	Q9zc27 yersinia pe
17	48	10.9	79	5	O17707	O17707 caenorhabdi
18	47.5	10.7	69	16	Q9EWV8	Q9ewv8 streptomyce
19	47.5	10.7	77	6	Q28593	Q28593 ovis aries
20	47.5	10.7	77	6	Q95MQ0	Q95mq0 ovis aries
21	47.5	10.7	77	16	Q8XCT7	Q8xct7 escherichia
22	47	10.6	45	2	Q52365	Q52365 escherichia
23	47	10.6	64	12	Q8UZB1	Q8uzb1 garlic late
24	47	10.6	77	16	Q932P0	Q932p0 staphylococ
25	46.5	10.5	55	4	Q8TBV0	Q8tbv0 homo sapien
26	46.5	10.5	72	6	Q9N072	Q9n072 macaca fasc
27	46	10.4	36	4	Q9BRI9	Q9bri9 homo sapien
28	46	10.4	39	5	Q9XYG8	Q9xyg8 plasmodium
29	46	10.4	57	10	Q8S1R9	Q8s1r9 oryza sativ
30	46	10.4	70	10	Q9FLN5	Q9fln5 arabidopsis
31	46	10.4	71	5	Q95NH5	Q95nh5 drosophila
32	46	10.4	80	8	Q35501	Q35501 paracentrot
33	45.5	10.3	48	4	Q9P1E0	Q9p1e0 homo sapien
34	45.5	10.3	60	5	Q9BJI5	Q9bji5 plasmodium
35	45.5	10.3	61	10	Q94DR7	Q94dr7 oryza sativ
36	45.5	10.3	66	10	Q943X6	Q943x6 oryza sativ
37	45.5	10.3	77	16	Q9KL85	Q9kl85 vibrio chol
38	45	10.2	32	8	Q36494	Q36494 farfantepen
39	45	10.2	38	5	P91965	P91965 penaeus van
40	45	10.2	40	8	Q9TGM1	Q9tgm1 homo sapien
41	45	10.2	44	11	Q04056	Q04056 rattus norv
42	45	10.2	49	2	Q50145	Q50145 mycobacteri
43	45	10.2	51	17	Q8TIN2	Q8tin2 methanosarc
44	45	10.2	54	16	Q92PQ7	Q92pq7 rhizobium m
45	45	10.2	56	13	O42594	O42594 xenopus lae

ALIGNMENTS

RESULT 1
Q96HV2
ID Q96HV2 PRELIMINARY; PRT; 57 AA.
AC Q96HV2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to glutamate receptor, metabotropic 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008050; AAH08050.1; -.
KW Receptor.
SQ SEQUENCE 57 AA; 5734 MW; DD4F557886DB753A CRC64;

Query Match 12.8%; Score 56.5; DB 4; Length 57;
Best Local Similarity 33.3%; Pred. No. 8.5;
Matches 18; Conservative 5; Mismatches 16; Indels 15; Gaps 3;

```
Qy      23 EGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPG 76
           :| | ||||| |  |        |||  |      : |:  : :|   |   ||
Db      13 KGGRGPAKAWEGGPTN------SSPNEAN------QQCEGTQGDP---LEAAPG 51
```

RESULT 2
Q8VZQ7
ID Q8VZQ7 PRELIMINARY; PRT; 78 AA.
AC Q8VZQ7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative defensin AMP1 protein.
GN AT1G19610.

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA   Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA   Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA   Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA   Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA   Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RT   "Full Length cDNA of gene At1g19610 (GI:15223595).";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY063933; AAL36289.1; -.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; UNKNOWN_1.
SQ   SEQUENCE   78 AA;  8840 MW;  A5B5DD28303A6545 CRC64;

  Query Match             12.8%;  Score 56.5;  DB 10;  Length 78;
  Best Local Similarity   36.5%;  Pred. No. 11;
  Matches   19;  Conservative    5;  Mismatches   13;  Indels   15;  Gaps    4;

Qy       7 SSLLCILLLCFSIF--------STEGK--RRPAKAWSG--RRTRLC---CHR 43
           ||   :| || |||         : ||:   | :| |:|    || |   | |
Db       3 SSYTLMLFLCLSIFLIASTEMMAVEGRICERRSKTWTGFCGNTRGCDSQCKR 54


RESULT 3
Q9BKX4
ID   Q9BKX4      PRELIMINARY;      PRT;    69 AA.
AC   Q9BKX4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 7.7 kDa protein.
GN   Y22D7AR.10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Madsen C.;
RT   "The sequence of C. elegans cosmid Y22D7AR.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC084154; AAK29877.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   69 AA;  7682 MW;  9DD66CD46976CF7E CRC64;

  Query Match             12.7%;  Score 56;  DB 5;  Length 69;
  Best Local Similarity   29.1%;  Pred. No. 12;
  Matches   23;  Conservative    8;  Mismatches   30;  Indels   18;  Gaps    6;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEGKRRPA---KAWSGRRTRLCCHRVPSPNSTNLKGHHV 57
           || ::  ||| :|| || :  |:          |: || |   ||     |:   |
Db       1 MRFIIFLSLL-LLLTCFKLGQTDTCDSYGCFPKSISGSR---CC-----PDGWEFGGR-- 49

Qy      58 RLCKPCKLEPE-PRLWVVP 75
             ||   |    :|  |
Db      50 ---APCPANPNMDEVWCCP 65


RESULT 4
Q9K1L9
ID   Q9K1L9      PRELIMINARY;      PRT;    75 AA.
AC   Q9K1L9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein NMB0092.
GN   NMB0092.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002368; AAF40554.1; -.
DR   TIGR; NMB0092; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   75 AA;  8504 MW;  B60F7F15CC8ACFB8 CRC64;

  Query Match             12.3%;  Score 54.5;  DB 16;  Length 75;
  Best Local Similarity   29.7%;  Pred. No. 20;
  Matches   22;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    4;

Qy       1 MRLLVLSSL----LCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCH-RVPSPNSTNLKGH 55
           :|  |||  |    || |  | | |            :|     |  | :| :||  |
Db       2 VRFFVLSFLTLINLCSLSACNSHF------------TGNINPLGTHNKVANPNCANSANS 49

Qy      56 HVRLCKPCKLEPEP 69
           |:|  :| :   :|
Db      50 HIR--QPSRKNYDP 61


RESULT 5
Q8UE52
ID   Q8UE52      PRELIMINARY;      PRT;    64 AA.
AC   Q8UE52;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein Atu1909.
GN   ATU1909.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
```

```
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
DR   EMBL; AE009145; AAL42905.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   64 AA;  7544 MW;  C15A5CA8687C42BE CRC64;

  Query Match             12.0%;  Score 53;  DB 16;  Length 64;
  Best Local Similarity   28.0%;  Pred. No. 26;
  Matches   14;  Conservative    6;  Mismatches   16;  Indels   14;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVV 74
           |:|||:                 | | | : ||||  : | :|  | :
Db      17 KKRPARR--------------PVKKTALSGGYRRLCKGMQWENDPTEWSI 52


RESULT 6
Q8VAU9
ID   Q8VAU9      PRELIMINARY;      PRT;    78 AA.
AC   Q8VAU9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wsv280 (WSSV335).
OS   White spot syndrome virus (WSSV).
OC   Viruses; unclassified viruses.
OX   NCBI_TaxID=92652;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21548311; PubMed=11689662;
RA   Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT   "Complete genome sequence of the shrimp white spot bacilliform
RT   virus.";
RL   J. Virol. 75:11811-11820(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RX   MEDLINE=20517548; PubMed=11062040;
RA   Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA   Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT   "Identification and characterization of a shrimp white spot syndrome
RT   virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT   cellular-type thymidine kinase and thymidylate kinase.";
RL   Virology 277:100-110(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RX   MEDLINE=21844071; PubMed=11853398;
RA   Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA   Lo C.F., Kou G.H.;
RT   "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT   spot syndrome virus and characterization of the motif important for
RT   targeting VP35 to the nuclei of transfected insect cells.";
RL   Virology 293:44-53(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RA   Lo C.-F., Kou G.-H.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF332093; AAL33283.1; -.
DR   EMBL; AF440570; AAL89203.1; -.
SQ   SEQUENCE   78 AA;  8949 MW;  B5800F9CFCB754D0 CRC64;

  Query Match             12.0%;  Score 53;  DB 12;  Length 78;
  Best Local Similarity   32.1%;  Pred. No. 32;
  Matches   18;  Conservative    7;  Mismatches   27;  Indels    4;  Gaps    2;

Qy       2 RLLVLSSLLCILLLCFSIFSTEGKRRPAKAWS-GRRTRLCCHRVPSPNSTNLKGHH 56
           | ||| : ||   :   || :    |  | ||   | :||  :  :|    : | |
Db      14 RQLVLFTSLCSSRVLELIFKSSSMLRRQKVWSLFHRAKLCTVQSGAP---MIAGQH 66


RESULT 7
Q8QN69
ID   Q8QN69      PRELIMINARY;      PRT;    69 AA.
AC   Q8QN69;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   EsV-1-221.
OS   Ectocarpus siliculosus virus.
OC   Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX   NCBI_TaxID=37665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ESV-1;
RA   Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RT   "The complete nucleotide sequence of the Ectocarpus siliculosus virus
RT   genome.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF204951; AAK14635.1; -.
SQ   SEQUENCE   69 AA;  8160 MW;  AACDBD15F312D639 CRC64;

  Query Match             11.9%;  Score 52.5;  DB 12;  Length 69;
  Best Local Similarity   25.8%;  Pred. No. 33;
  Matches   16;  Conservative   10;  Mismatches   23;  Indels   13;  Gaps    3;

Qy      19 IFSTEG---------KRRPAKAWSGR---RTRLCCHRVPSPNS-TNLKGHHVRLCKPCKL 65
           |||  |         : |    ||      | |: |:|: |  :  : :| | ::|    :
Db       2 IFSVNGILNTKLRLRRHRLCHLWSNNLLPRRRIPCNRLQSRQARVSCRGAHKKICTGGYI 61

Qy      66 EP 67
           |
Db      62 RP 63


RESULT 8
Q985E2
ID   Q985E2      PRELIMINARY;      PRT;    51 AA.
AC   Q985E2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein msl7717.
GN   MSL7717.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003012; BAB54120.1; -.
```

```
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   51 AA;  5843 MW;  4934C9A7166468F7 CRC64;

  Query Match           11.5%;  Score 51;  DB 16;  Length 51;
  Best Local Similarity  32.0%;  Pred. No. 38;
  Matches   16;  Conservative   12;  Mismatches   12;  Indels   10;  Gaps    5;

Qy       8 SLLCILL----LCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 53
           :::|:|:    : |:|   |  | | ::|    ||:|  |:  |  |||:
Db       7 TIVCLLMSLFGMFFAILVAEAGRSP-RSWD-EATRIC--RLACP--TNLR 50


RESULT 9
Q9KKN6
ID   Q9KKN6      PRELIMINARY;      PRT;    79 AA.
AC   Q9KKN6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein VCA1066.
GN   VCA1066.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004432; AAF96960.1; -.
DR   TIGR; VCA1066; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   79 AA;  8342 MW;  716F1729FE8C41C2 CRC64;

  Query Match           11.5%;  Score 51;  DB 16;  Length 79;
  Best Local Similarity  42.9%;  Pred. No. 58;
  Matches    9;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      39 LCCHRVPSPNSTNLKGHHVRL 59
           | ||   :|:|  : | ||:|
Db      50 LTCHLPATPSSLGIAGSHVKL 70


RESULT 10
Q32989
ID   Q32989      PRELIMINARY;      PRT;    42 AA.
AC   Q32989;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORF42f.
OS   Pinus thunbergii (Green pine) (Japanese black pine).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=3350;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92212283; PubMed=1557027;
RA   Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,
RA   Wakasugi T., Sugiura M.;
RT   "Chloroplast DNA of black pine retains a residual inverted repeat
RT   lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
RT   trnH and the absence of rps16.";
RL   Mol. Gen. Genet. 232:206-214(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95094312; PubMed=8001170;
RA   Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiura M.;
RT   "A new gene encoding tRNA pro (GGG) is present in the chloroplast
RT   genome of black pine: a compilation of 32 tRNA genes from black pine
RT   chloroplasts.";
RL   Curr. Genet. 26:153-158(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95024047; PubMed=7937893;
RA   Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA   Sugiura M.;
RT   "Loss of all ndh genes as determined by sequencing the entire
RT   chloroplast genome of the black pine Pinus thunbergii.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
DR   EMBL; D17510; BAA04433.1; -.
KW   Chloroplast.
SQ   SEQUENCE   42 AA;  4480 MW;  0B5B2669DB4AD8A4 CRC64;

  Query Match           11.3%;  Score 50;  DB 8;  Length 42;
  Best Local Similarity  37.5%;  Pred. No. 42;
  Matches    6;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      63 CKLEPEPRLWVVPGAL 78
           |   |:| :|:: | |
Db       6 CSSAPDPEMWIIQGTL 21


RESULT 11
Q9Q6Q2
ID   Q9Q6Q2      PRELIMINARY;      PRT;    59 AA.
AC   Q9Q6Q2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Grapevine leafroll-associated virus 1 helicase (HEL) and RNA-dependent
DE   RNA polymerase (POL) genes; and p7, HSP70-like protein, p55, coat
DE   protein (CP), p55 (CPd1), p50 (CPd2), p22, and p24 genes.
OS   grapevine leafroll-associated virus 1.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC   Closterovirus.
OX   NCBI_TaxID=47985;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20141373; PubMed=10675398;
RA   Fazeli C.F., Rezaian M.A.;
RT   "Nucleotide sequence and organization of ten open reading frames in
RT   the genome of Grapevine leafroll-associated virus 1 and identification
RT   of three subgenomic RNAs.";
RL   J. Gen. Virol. 81:605-615(2000).
DR   EMBL; AF195822; AAF22739.1; -.
SQ   SEQUENCE   59 AA;  6736 MW;  B90BED563F81DCF2 CRC64;

  Query Match           11.3%;  Score 50;  DB 12;  Length 59;
  Best Local Similarity  25.9%;  Pred. No. 59;
  Matches   15;  Conservative   10;  Mismatches   19;  Indels   14;  Gaps    2;

Qy       3 LLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRV---PSPNSTNLKGHHV 57
           ||   ||  :::|||:::           :  |  | || ::   |    ||.   :|
Db      10 LLYTVSLFVLVVLCFAVY----------FVVRAIRHCCAKIEDKPVGGPTNSSFRYV 56


RESULT 12
Q91KH1
ID   Q91KH1      PRELIMINARY;      PRT;    76 AA.
AC   Q91KH1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
```

```
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H33;
RA   Gimenez-Barcons M., Franco S., Suarez Y., Forns X., Ampurdanes S.,
RA   Puig-Basagoiti F., Sanchez-Fueyo A., Barrera J.M., Llovet J.M.,
RA   Bruix J., Sanchez-Tapias J.M., Rodes J., Saiz J.C.;
RT   "High amino acid variability within the NS5A gene of hepatitis C virus
RT   (HCV) is associated with hepatocellular carcinoma in patients with
RT   HCV-1b related cirrhosis.";
RL   Hepatology 0:0-0(2001).
DR   EMBL; AF379099; AAK63318.1; -.
DR   InterPro; IPR002522; HCV_capsid.
DR   Pfam; PF01543; HCV_capsid; 1.
FT   NON_TER       1      1
FT   NON_TER      76     76
SQ   SEQUENCE   76 AA;  8660 MW;  394FC42024A5B4A8 CRC64;

  Query Match             11.3%;  Score 50;  DB 12;  Length 76;
  Best Local Similarity   33.3%;  Pred. No. 75;
  Matches   14;  Conservative    4;  Mismatches   16;  Indels    8;  Gaps    2;

Qy      35 RRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPG 76
           | ||    |      :| :|    : | |  :|| | |  ||
Db      36 RATRKTSER------SNPRGRRPPISKAC--QPERRTWAHPG 69


RESULT 13
Q9JHY4
ID   Q9JHY4      PRELIMINARY;      PRT;    80 AA.
AC   Q9JHY4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Elafin-like protein I (SWAM1 protein).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ;
RA   Hagiwara K., Hoshi S., Takahashi M., Miki T., Nukiwa T.;
RT   "ELM-like proteins.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hagiwara K., Kikuchi T., Endo Y., Huqun, Takahashi M., Xin X.,
RA   Hoshi S., Miki M., Inooka N., Tokue Y., Nukiwa T.;
RT   "Mouse SWAM1 (single WAP motif protein 1) gene.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF276974; AAF86471.1; -.
DR   EMBL; AF482009; AAL90747.1; -.
DR   HSSP; O46655; 1CJH.
DR   InterPro; IPR002221; WAP.
DR   Pfam; PF00095; wap; 1.
DR   PRINTS; PR00003; 4DISULPHCORE.
DR   SMART; SM00217; WAP; 1.
DR   PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
SQ   SEQUENCE   80 AA;  9237 MW;  9135647ED91F63ED CRC64;

  Query Match             11.2%;  Score 49.5;  DB 11;  Length 80;
  Best Local Similarity   34.7%;  Pred. No. 92;
  Matches   17;  Conservative    8;  Mismatches   17;  Indels    7;  Gaps    2;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCC--HRVPSP 47
           |:|| || |   :||| ::   | |::   :  |      |  :||| |
Db       1 MKLLGLSLLAVTILLCCNMARPEIKKKNVFSKPG-----YCPEYRVPCP 44


RESULT 14
```

```
Q9VU66
ID   Q9VU66      PRELIMINARY;      PRT;    71 AA.
AC   Q9VU66;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG10714 protein.
GN   LY OR CG10714.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003538; -; NOT_ANNOTATED_CDS.
DR   FlyBase; FBgn0002573; Ly.
SQ   SEQUENCE   71 AA;  7887 MW;  D5AEF002A90AFC47 CRC64;

  Query Match             11.0%;  Score 48.5;  DB 5;  Length 71;
  Best Local Similarity   29.6%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches   15;  Indels   23;  Gaps    2;

Qy      30 KAWSGRR----------TRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWV 73
           | | |           | |  |||              | |    | |||||
Db      29 KQWMQRASAFNTICPFCTTRCSTRVP------------YCGPNFCCPRPRLWV 69


RESULT 15
Q61910
ID   Q61910      PRELIMINARY;      PRT;    61 AA.
AC   Q61910;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Mammary transforming protein.
GN   PEA15 OR MAT1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=BREAST TUMOR;
RX   MEDLINE=95024046; PubMed=7937892;
RA   Bera T.K., Guzman R.C., Miyamoto S., Panda D.K., Sasaki M., Hanyu K.,
RA   Enami J., Nandi S.;
RT   "Identification of a mammary transforming gene (MAT1) associated with
RT   mouse mammary carcinogenesis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:9789-9793(1994).
DR   EMBL; L31958; AAC37665.1; -.
DR   MGD; MGI:104799; Pea15.
SQ   SEQUENCE   61 AA;  7137 MW;  DDC1F91842892171 CRC64;

  Query Match             10.9%;  Score 48;  DB 11;  Length 61;
  Best Local Similarity   28.3%;  Pred. No. 1.1e+02;
  Matches   13;  Conservative    7;  Mismatches   16;  Indels   10;  Gaps    2;

Qy      17 FSIFSTEGKRRPAKAWSGR------RTRLCCHRVPSPNSTNLKGHH 56
           | :||     || : | |       | :|    : |||  ::  :|
Db      13 FVVFSINLLSRPEREWEGMPQKGSGRAKL----LQSPNRKHISTNH 54


Search completed: May  8, 2003, 21:47:37
Job time : 33 secs
```

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:42:53 ; Search time 10 Seconds
(without alignments)
335.958 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 7253

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	59	13.3	55	1	ATP8_LATCH	O03168	latimeria c
2	53	12.0	56	1	RS14_PYRHO	O74093	pyrococcus
3	48.5	11.0	78	1	RT19_PROWI	P46750	prototheca
4	48	10.9	61	1	R14A_LACLA	Q9cdx5	lactococcus
5	48	10.9	74	1	SGS7_DROME	P02841	drosophila
6	47	10.6	64	1	BD01_SHEEP	O19038	ovis aries
7	47	10.6	69	1	RS7_METMT	O93636	methanococc
8	46.5	10.5	69	1	BD01_MOUSE	P56386	mus musculu
9	46	10.4	69	1	BD01_RAT	O89117	rattus norv
10	45	10.2	75	1	SGS8_DROME	P02842	drosophila
11	44.5	10.1	46	1	AX2_BETVU	P82010	beta vulgar
12	44.5	10.1	61	1	R14A_STRPY	Q9alw1	streptococc
13	44.5	10.1	75	1	VGE_BPPHK	Q38040	bacteriopha
14	44	10.0	50	1	HSP1_GORGO	P35303	gorilla gor
15	44	10.0	59	1	SCK2_ANDMA	P45696	androctonus
16	44	10.0	63	1	ITR4_LUFCY	P34950	luffa cylin
17	43.5	9.8	75	1	VGE_BPAL3	P31280	bacteriopha
18	43	9.7	49	1	HSP1_RABIT	P10119	oryctolagus
19	43	9.7	63	1	FCA1_CRIFA	Q23688	crithidia f
20	43	9.7	68	1	L2MU_ADECC	Q65953	canine aden
21	43	9.7	75	1	YCEK_ECOLI	P45806	escherichia
22	42.5	9.6	54	1	RL37_THEVO	Q97bu6	thermoplasm
23	42.5	9.6	55	1	NXB4_CERLA	P01525	cerebratulu
24	42.5	9.6	57	1	RL32_STAAM	Q99ux6	staphylococ
25	42.5	9.6	66	1	CXA2_CONTE	Q9xzk7	conus texti
26	42.5	9.6	69	1	CXA1_CONTE	Q9xzk6	conus texti
27	42.5	9.6	72	1	VG18_BPMU	Q38625	bacteriopha
28	42.5	9.6	80	1	AFP2_RAPSA	P30230	raphanus sa
29	42	9.5	49	1	HSP1_SAGIM	P24714	saguinus im
30	42	9.5	50	1	HSP1_BOVIN	P02318	bos taurus
31	42	9.5	50	1	HSP1_SHEEP	P04102	ovis aries
32	42	9.5	50	1	PND1_ECOLI	P11902	escherichia
33	42	9.5	70	1	YVAK_VACCC	P20520	vaccinia vi
34	41.5	9.4	53	1	RS14_METVA	P14041	methanococc
35	41.5	9.4	64	1	LAP_BOVIN	Q28880	bos taurus
36	41	9.3	27	1	CXM3_CONPU	P58925	conus purpu
37	41	9.3	52	1	V07K_FXMV	P22171	foxtail mos
38	41	9.3	63	1	RL35_CAMJE	Q9piq1	campylobact
39	41	9.3	70	1	RS21_RALSO	Q8y3h2	ralstonia s
40	41	9.3	72	1	NXL1_DENJA	P01393	dendroaspis
41	41	9.3	79	1	FRD1_AQUAE	O67251	aquifex aeo
42	40.5	9.2	54	1	RL37_THEAC	Q9hit7	thermoplasm
43	40.5	9.2	61	1	RZ1_BP933	Q9xjj7	bacteriopha
44	40.5	9.2	77	1	LEA2_MACMU	Q95m25	macaca mula
45	40.5	9.2	77	1	Y125_HUMAN	Q14138	homo sapien

ALIGNMENTS

RESULT 1
ATP8_LATCH
ID ATP8_LATCH STANDARD; PRT; 55 AA.
AC O03168;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN MTATP8 OR ATP8.
OS Latimeria chalumnae (Latimeria) (Coelacanth).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OX NCBI_TaxID=7897;
RN [1]
RP SEQUENCE FROM N.A.
RA Zardoya R., Meyer A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC H(+)(Out).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC ---

CC ---
DR EMBL; U82228; AAC60322.1; -.
DR InterPro; IPR001421; ATPase8_mit.
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 55 AA; 6569 MW; 7FE36319E8AF825B CRC64;

Query Match 13.3%; Score 59; DB 1; Length 55;
Best Local Similarity 34.3%; Pred. No. 3.9;
Matches 23; Conservative 7; Mismatches 15; Indels 22; Gaps 5;

```
Qy       7 SSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPS-PNSTNLKGHHVRLCKPCKL 65
           |  | |||  : || |    |:|      |:|  |  |: |:: |:          ||
Db       7 SPWLLILLFSWLIFLT---MLPSK------TQL--HTFPNMPSTQNM----------CKQ 45

Qy      66 EPEPRLW 72
           ||||  |
Db      46 EPEPWTW 52
```

RESULT 2
RS14_PYRHO
ID RS14_PYRHO STANDARD; PRT; 56 AA.

```
AC   O74093;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   30S ribosomal protein S14P.
GN   RPS14P OR PHS047 OR PAB7080.
OS   Pyrococcus horikoshii, and
OS   Pyrococcus abyssi.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=53953, 29292;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=P.horikoshii; STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper-
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=P.abyssi; STRAIN=GE5 / Orsay;
RA   Heilig R.;
RT   "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT   structure and evolution.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP000007; BAA30879.1; -.
DR   EMBL; AJ248284; CAB49249.1; -.
DR   InterPro; IPR001209; Ribosomal_S14.
DR   Pfam; PF00253; Ribosomal_S14; 1.
DR   PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   56 AA;  6626 MW;  B5D97E9E2841F686 CRC64;

  Query Match            12.0%;  Score 53;  DB 1;  Length 56;
  Best Local Similarity  29.8%;  Pred. No. 18;
  Matches   14;  Conservative    8;  Mismatches   21;  Indels    4;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRL 71
           ||:| |   | |  : | :        :: | : ||: |  |  |:|
Db       8 KRKPRKFGKGARRCIRCGQY----GPIIRIHGLMLCRHCFREVAPKL 50


RESULT 3
RT19_PROWI
ID   RT19_PROWI     STANDARD;      PRT;    78 AA.
AC   P46750;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Mitochondrial ribosomal protein S19.
GN   RPS19.
OS   Prototheca wickerhamii.
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellaceae; Prototheca.
OX   NCBI_TaxID=3111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=263-11;
RX   MEDLINE=94180393; PubMed=8133522;
RA   Wolff G., Plante I., Lang B.F., Kueck U., Burger G.;
RT   "Complete sequence of the mitochondrial DNA of the chlorophyte alga
RT   Prototheca wickerhamii. Gene content and genome organization.";
RL   J. Mol. Biol. 237:75-86(1994).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U02970; AAD12637.1; -.
DR   HSSP; P80381; 1QKF.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   Pfam; PF00203; Ribosomal_S19; 1.
DR   PRINTS; PR00975; RIBOSOMALS19.
DR   ProDom; PD001012; Ribosomal_S19; 1.
DR   TIGRFAMs; TIGR01050; rpsS_bact; 1.
DR   PROSITE; PS00323; RIBOSOMAL_S19; 1.
KW   Ribosomal protein; Mitochondrion.
SQ   SEQUENCE   78 AA;  8847 MW;  9B53BDA9A287BCDB CRC64;

  Query Match            11.0%;  Score 48.5;  DB 1;  Length 78;
  Best Local Similarity  33.3%;  Pred. No. 77;
  Matches   13;  Conservative    7;  Mismatches   14;  Indels    5;  Gaps    2;

Qy      30 KAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPE 68
           | || ||: :    :|      :  |: |:  |||: ||
Db      23 KIWS-RRSAI----LPQFVGKTVSIHNGRIFIPCKISPE 56


RESULT 4
R14A_LACLA
ID   R14A_LACLA     STANDARD;      PRT;    61 AA.
AC   Q9CDX5;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S14-1.
GN   RPSN1 OR LL2086.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
CC   -!- FUNCTION: Known to be required for the assembly of 30S particles
CC       and may also be responsible for determining the conformation of
CC       the 16S rRNA at the A site (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006437; AAK06184.1; -.
```

```
DR   InterPro; IPR001209; Ribosomal_S14.
DR   Pfam; PF00253; Ribosomal_S14; 1.
DR   PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   61 AA;  7135 MW;  73101824196457B4 CRC64;

  Query Match            10.9%;  Score 48;  DB 1;  Length 61;
  Best Local Similarity  33.9%;  Pred. No. 70;
  Matches   19;  Conservative    5;  Mismatches   22;  Indels   10;  Gaps    3;

Qy      26 RRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 81
           :|||| :   || | |   |:|   |    :||: | |         | || |
Db      11 QRPAKFSTQAYTR--CERCGRPHSVYRK---FKLCRICLRE-----LAYKGQLPGV 56


RESULT 5
SGS7_DROME
ID   SGS7_DROME     STANDARD;      PRT;     74 AA.
AC   P02841; Q9VTJ0;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Salivary glue protein Sgs-7 precursor.
GN   SGS7 OR CG18087.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83294545; PubMed=6411930;
RA   Garfinkel M.D., Pruitt R.E., Meyerowitz E.M.;
RT   "DNA sequences, gene regulation and modular protein evolution in the
RT   Drosophila 68C glue gene cluster.";
RL   J. Mol. Biol. 168:765-789(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- DEVELOPMENTAL STAGE: PRODUCED BY THIRD-INSTAR LARVAE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01918; CAA25993.1; -.
DR   EMBL; AE003544; AAF50058.1; -.
DR   PIR; A03330; GSFF7.
DR   FlyBase; FBgn0003377; Sgs7.
KW   Signal.
FT   SIGNAL        1     23
FT   CHAIN        24     74       SALIVARY GLUE PROTEIN SGS-7.
SQ   SEQUENCE   74 AA;  7919 MW;  914ACA5DB9153E29 CRC64;

  Query Match            10.9%;  Score 48;  DB 1;  Length 74;
  Best Local Similarity  26.8%;  Pred. No. 84;
  Matches   22;  Conservative   18;  Mismatches   26;  Indels   16;  Gaps    7;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEG--KRRP----AKAWSG--RRTRLCCHRVPSPNSTNL 52
           |:|: :: : ||||: ||  :  |  : :|     || :|   : :||   :    :
Db       1 MKLIAVTIIACILLIGFSDLALGGACECQPCGPGGKACTGCPEKPQLCQQLISDIRNLQQ 60

Qy      53 KGHHVRLCKPCKLEPEPRLWVV 74
           |   :| |  |  ||:   |::
Db      61 K---IRKC-VCG-EPQ---WMI 74


RESULT 6
BD01_SHEEP
ID   BD01_SHEEP     STANDARD;      PRT;     64 AA.
AC   O19038;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-defensin 1 precursor (BD-1) (sBD1).
GN   DEFB1.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98138497; PubMed=9478010;
RA   Huttner K.M., Brezinski-Caliguri D.J., Mahoney M.M., Diamond G.;
RT   "Antimicrobial peptide expression is developmentally regulated in the
RT   ovine gastrointestinal tract.";
RL   J. Nutr. 128:297S-299S(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Trachea;
RX   MEDLINE=98121317; PubMed=9461419;
RA   Huttner K.M., Lambeth M.R., Burkin H.R., Broad T.E.;
RT   "Localization and genomic organization of sheep antimicrobial peptides
RT   genes.";
RL   Gene 206:85-91(1998).
CC   -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U75250; AAB61995.1; -.
DR   HSSP; P46170; 1BNB.
DR   InterPro; IPR001855; Defensin_beta.
DR   InterPro; IPR001271; Defensin_mammal.
DR   Pfam; PF00711; Defensin_beta; 1.
DR   SMART; SM00048; DEFSN; 1.
KW   Antibiotic; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23     64       BETA-DEFENSIN 1.
FT   DISULFID     31     60       BY SIMILARITY.
FT   DISULFID     38     53       BY SIMILARITY.
FT   DISULFID     43     61       BY SIMILARITY.
SQ   SEQUENCE   64 AA;  7244 MW;  3529A9B76ABD023A CRC64;

  Query Match             10.6%;  Score 47;  DB 1;  Length 64;
  Best Local Similarity   30.2%;  Pred. No. 95;
  Matches   19;  Conservative    7;  Mismatches   19;  Indels   18;  Gaps    4;

Qy      12 ILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHR-----VPS--PNSTNLKGHHVRLCKPCK 64
           :||: | :  : |    :    | | || |||     |||  |        |:|    |:
Db       6 LLLVLFFVVLSAG----SGFTQGVRNRLSCHRNKGVCVPSRCP-------RHMRQIGTCR 54

Qy      65 LEP 67
             |
Db      55 GPP 57


RESULT 7
RS7_METMT
ID   RS7_METMT      STANDARD;      PRT;    69 AA.
AC   O93636;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   30S ribosomal protein S7P (Fragment).
GN   RPS7P OR S7.
OS   Methanococcoides methylutens.
OC   Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC   Methanosarcinaceae; Methanococcoides.
OX   NCBI_TaxID=2226;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 2657T;
RX   MEDLINE=99059471; PubMed=9845338;
RA   Thomas T., Cavicchioli R.;
RT   "Archaeal cold-adapted proteins: structural and evolutionary analysis
RT   of the elongation factor 2 proteins from psychrophilic, mesophilic and
RT   thermophilic methanogens.";
RL   FEBS Lett. 439:281-287(1998).
CC   -!- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
CC       16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF022779; AAC79196.1; -.
DR   InterPro; IPR000235; Ribosomal_S7.
DR   Pfam; PF00177; Ribosomal_S7; 1.
DR   PROSITE; PS00052; RIBOSOMAL_S7; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER       1      1
SQ   SEQUENCE   69 AA;  7488 MW;  7E64F12BDA024160 CRC64;

  Query Match             10.6%;  Score 47;  DB 1;  Length 69;
  Best Local Similarity   62.5%;  Pred. No. 1e+02;
  Matches   10;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 CFSIFSTEGKRRPAKA 31
           ||||   :|| | |||
Db      52 CFSINRKDGKERVAKA 67


RESULT 8
BD01_MOUSE
ID   BD01_MOUSE     STANDARD;      PRT;    69 AA.
AC   P56386;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Beta-defensin 1 precursor (BD-1) (mBD-1).
GN   DEFB1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
RA   Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
RA   Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
RA   Theising B., Wylie T., Lennon G., Soares B., Wilson R., Waterston R.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97431609; PubMed=9287114;
RA   Huttner K.M., Kozak C.A., Bevins C.L.;
RT   "The mouse genome encodes a single homolog of the antimicrobial
RT   peptide human beta-defensin 1.";
RL   FEBS Lett. 413:45-49(1997).
RN   [3]
RP   TISSUE SPECIFICITY.
RC   STRAIN=C57BL/6, 129/SvJ, and FVB; TISSUE=Lung;
RX   MEDLINE=20517883; PubMed=10922379;
RA   Jia H.P., Wowk S.A., Schutte B.C., Lee S.K., Vivado A., Tack B.F.,
RA   Bevins C.L., McCray P.B. Jr.;
RT   "A novel murine beta-defensin expressed in tongue, esophagus, and
RT   trachea.";
RL   J. Biol. Chem. 275:33314-33320(2000).
CC   -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: DETECTED IN KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AA071757; -; NOT_ANNOTATED_CDS.
DR   EMBL; AA065510; -; NOT_ANNOTATED_CDS.
DR   EMBL; AA108061; -; NOT_ANNOTATED_CDS.
DR   EMBL; AA107538; -; NOT_ANNOTATED_CDS.
DR   EMBL; AA105324; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF003525; AAB72003.1; -.
DR   EMBL; AF003524; AAB72003.1; JOINED.
DR   MGD; MGI:1096878; Defb1.
DR   InterPro; IPR001855; Defensin_beta.
DR   InterPro; IPR001271; Defensin_mammal.
DR   Pfam; PF00711; Defensin_beta; 1.
DR   SMART; SM00048; DEFSN; 1.
KW   Antibiotic; Signal.
FT   SIGNAL        1     21       POTENTIAL.
```

```
FT   PROPEP       22      32       BY SIMILARITY.
FT   CHAIN        33      69       BETA-DEFENSIN 1.
FT   DISULFID     37      66       BY SIMILARITY.
FT   DISULFID     44      59       BY SIMILARITY.
FT   DISULFID     49      67       BY SIMILARITY.
SQ   SEQUENCE   69 AA;  7749 MW;  8B12DBD3BBF3ACE4 CRC64;

  Query Match             10.5%;  Score 46.5;  DB 1;  Length 69;
  Best Local Similarity   31.1%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    8;  Mismatches   21;  Indels   13;  Gaps    4;

Qy      12 ILLLCFSIFSTEGKRRPAKAWSGRRT--RLC------CHRVPSPNSTNLKGHHVRLCKPC 63
           ::::|| :||           ||||    |       | |   |::| |:|     |||
Db       8 LVMICF-LFSQMEPGVGILTSLGRRTDQYKCLQHGGFCLRSSCPSNTKLQG----TCKPD 62

Qy      64 K 64
           |
Db      63 K 63


RESULT 9
BD01_RAT
ID   BD01_RAT       STANDARD;      PRT;    69 AA.
AC   O89117;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-defensin 1 precursor (BD-1) (RBD-1).
GN   DEFB1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Goto-Kakizaki; TISSUE=Kidney;
RA   Page R.A., Malik A.N.;
RT   "Rat beta defensin-1 peptide: candidate marker for diabetic
RT   nephropathy.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=99386883; PubMed=10456937;
RA   Jia H.P., Mills J.N., Barahmand-Pour F., Nishimura D.,
RA   Mallampali R.K., Wang G., Wiles K., Tack B.F., Bevins C.L.,
RA   McCray P.B. Jr.;
RT   "Molecular cloning and characterization of rat genes encoding
RT   homologues of human beta-defensins.";
RL   Infect. Immun. 67:4827-4833(1999).
CC   -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Highly expressed in kidney.
CC   -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF093536; AAC61871.1; -.
DR   EMBL; AF068860; AAC28071.1; -.
DR   InterPro; IPR001855; Defensin_beta.
DR   InterPro; IPR001271; Defensin_mammal.
DR   Pfam; PF00711; Defensin_beta; 1.
DR   SMART; SM00048; DEFSN; 1.
KW   Antibiotic; Signal.
FT   SIGNAL        1      21       POTENTIAL.
FT   PROPEP       22      32       BY SIMILARITY.
FT   CHAIN        33      69       BETA-DEFENSIN 1.
FT   DISULFID     37      66       BY SIMILARITY.
FT   DISULFID     44      59       BY SIMILARITY.
FT   DISULFID     49      67       BY SIMILARITY.
SQ   SEQUENCE   69 AA;  7837 MW;  66B1F0C29BB5C991 CRC64;

  Query Match             10.4%;  Score 46;  DB 1;  Length 69;
  Best Local Similarity   31.1%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    6;  Mismatches   24;  Indels   12;  Gaps    3;

Qy      12 ILLLCFSIFSTEGKRRPAKAWSGRRT--RLC------CHRVPSPNSTNLKGHHVRLCKPC 63
           :|:: | :||           ||||    |       | |   |: | |:|     |||
Db       7 LLVMLFFLFSQMELGAGILTSLGRRTDQYRCLQNGGFCLRSSCPSHTKLQG----TCKPD 62

Qy      64 K 64
           |
Db      63 K 63


RESULT 10
SGS8_DROME
ID   SGS8_DROME     STANDARD;      PRT;    75 AA.
AC   P02842; Q9VTI9;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Salivary glue protein Sgs-8 precursor.
GN   SGS8 OR CG6132.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83294545; PubMed=6411930;
RA   Garfinkel M.D., Pruitt R.E., Meyerowitz E.M.;
RT   "DNA sequences, gene regulation and modular protein evolution in the
RT   Drosophila 68C glue gene cluster.";
RL   J. Mol. Biol. 168:765-789(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
```

```
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- DEVELOPMENTAL STAGE: PRODUCED BY THIRD-INSTAR LARVAE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01918; CAA25992.1; -.
DR   EMBL; AE003544; AAF50059.1; -.
DR   PIR; A03331; GSFF8.
DR   FlyBase; FBgn0003378; Sgs8.
KW   Signal.
FT   SIGNAL       1     24
FT   CHAIN       25     75       SALIVARY GLUE PROTEIN SGS-8.
SQ   SEQUENCE   75 AA;  7917 MW;  27B818DFE158C636 CRC64;

  Query Match             10.2%;  Score 45;  DB 1;  Length 75;
  Best Local Similarity   44.4%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MRLLVLSSLLCILLLCFS 18
           |:|||:: : ||:|: |:
Db       1 MKLLVVAVIACIMLIGFA 18


RESULT 11
AX2_BETVU
ID   AX2_BETVU      STANDARD;      PRT;    46 AA.
AC   P82010; P81510;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Antifungal protein AX2.
OS   Beta vulgaris (Sugar beet).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX   NCBI_TaxID=3555;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. Turbo, and cv. Rhizor; TISSUE=Leaf;
RX   MEDLINE=95383713; PubMed=7655063;
RA   Kragh K.M., Nielsen J.E., Nielsen K.K., Dreboldt S., Mikkelsen J.D.;
RT   "Characterization and localization of new antifungal cysteine-rich
RT   proteins from Beta vulgaris.";
RL   Mol. Plant Microbe Interact. 8:424-434(1995).
CC   -!- FUNCTION: STRONG INHIBITING ACTIVITY AGAINST C.BETICOLA AND OTHER
CC       FILAMENTOUS FUNGI. LITTLE OR NO EFFECT AGAINST BACTERIA.
CC   -!- TISSUE SPECIFICITY: LEAVES AND FLOWERS.
CC   -!- MASS SPECTROMETRY: MW=5193; MW_ERR=3; METHOD=Electrospray.
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
DR   HSSP; P20158; 1GPS.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knot1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; Gamma-thionin; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Plant defense; Antibiotic; Fungicide.
FT   DISULFID      3     46       BY SIMILARITY.
FT   DISULFID     14     34       BY SIMILARITY.
FT   DISULFID     20     40       BY SIMILARITY.
FT   DISULFID     24     42       BY SIMILARITY.
SQ   SEQUENCE   46 AA;  5185 MW;  9A536EE9E74B18A6 CRC64;

  Query Match             10.1%;  Score 44.5;  DB 1;  Length 46;
  Best Local Similarity   27.9%;  Pred. No. 1.3e+02;
  Matches   12;  Conservative    6;  Mismatches   20;  Indels    5;  Gaps    1;

Qy      26 RRPAKAWSG-----RRTRLCCHRVPSPNSTNLKGHHVRLCKPC 63
             |:|:  :||         :  |:|   ||   | |      :||
Db       4 RKPSMYFSGACFSDTNCQKACNREDWPNGKCLVGFKCECQRPC 46


RESULT 12
R14A_STRPY
ID   R14A_STRPY     STANDARD;      PRT;    61 AA.
AC   Q9A1W1;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S14-1.
GN   RPSN1 OR SPY0064 OR SPYM18_0065.
OS   Streptococcus pyogenes, and
OS   Streptococcus pyogenes (serotype M18).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314, 186103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS8232 / Serotype M18;
RX   MEDLINE=21927593; PubMed=11917108;
RA   Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA   Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA   Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA   Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT   "Genome sequence and comparative microarray analysis of serotype M18
RT   group A Streptococcus strains associated with acute rheumatic fever
RT   outbreaks.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC   -!- FUNCTION: Known to be required for the assembly of 30S particles
CC       and may also be responsible for determining the conformation of
CC       the 16S rRNA at the A site (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006478; AAK33195.1; -.
DR   EMBL; AE009959; AAL96889.1; -.
DR   InterPro; IPR001209; Ribosomal_S14.
DR   Pfam; PF00253; Ribosomal_S14; 1.
DR   PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   61 AA;  7073 MW;  3AE47B92FD4840BF CRC64;

  Query Match             10.1%;  Score 44.5;  DB 1;  Length 61;
  Best Local Similarity   34.2%;  Pred. No. 1.7e+02;
  Matches   13;  Conservative    6;  Mismatches   14;  Indels    5;  Gaps    2;

Qy      26 RRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPC 63
```

```
          :|||| :  || | :  |:|  |   :||: |
Db     11 KRPAKHSTQAYTR--CEKCGRPHSVYRK---FKLCRVC 43


RESULT 13
VGE_BPPHK
ID   VGE_BPPHK      STANDARD;      PRT;    75 AA.
AC   Q38040;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Lysis protein (E protein) (GPE).
GN   E.
OS   Bacteriophage phi-K.
OC   Viruses; ssDNA viruses; Microviridae; Microvirus.
OX   NCBI_TaxID=10848;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: E PROTEIN IS RESPONSIBLE FOR HOST CELL LYSIS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60323; CAA42889.1; -.
KW   Phage lysis protein.
SQ   SEQUENCE   75 AA;  8441 MW;  7F900ECDC8A59045 CRC64;

  Query Match          10.1%;  Score 44.5;  DB 1;  Length 75;
  Best Local Similarity  30.0%;  Pred. No. 2.1e+02;
  Matches   21;  Conservative    5;  Mismatches   33;  Indels   11;  Gaps    3;

Qy      3 LLVLSSLLCILLLCF--SIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLC 60
          ||:|| ||  ||: |  | |     |: |  |  |          |   | |
Db     13 LLLLSLLLPSLLIMFIPSTF-----RQHASLWKARSLAKTLLMASSARLTPLSSSRT--- 64

Qy     61 KPCKLEPEPR 70
            || |  : :
Db     65 -PCVLRQDSK 73


RESULT 14
HSP1_GORGO
ID   HSP1_GORGO     STANDARD;      PRT;    50 AA.
AC   P35303;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Sperm protamine P1 (Cysteine-rich protamine).
GN   PRM1.
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94141943; PubMed=8308910;
RA   Retief J.D., Winkfein R.J., Dixon G.H., Adroer R., Queralt R.,
RA   Ballabriga J., Oliva R.;
RT   "Evolution of protamine P1 genes in primates.";
RL   J. Mol. Evol. 37:426-434(1993).
CC   -!- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC       SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC       SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC   -!- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
CC       DNA-HELIX (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14587; AAA51528.1; -.
DR   InterPro; IPR000221; Protamine_P1.
DR   Pfam; PF00260; protamine_P1; 1.
DR   PROSITE; PS00048; PROTAMINE_P1; 1.
KW   Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW   Testis; DNA condensation; Nuclear protein.
FT   INIT_MET      0      0      BY SIMILARITY.
SQ   SEQUENCE   50 AA;  6600 MW;  0EE9ABB8451EFEA2 CRC64;

  Query Match          10.0%;  Score 44;  DB 1;  Length 50;
  Best Local Similarity  32.1%;  Pred. No. 1.6e+02;
  Matches    9;  Conservative    3;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     16 CFSIFSTEGKRRPAKAWSGRRTRLCCHR 43
          |:    |  :||     : ||   || |
Db     14 CYRQRQTSRRRRRSCQTQRRAMRCCRR 41


RESULT 15
SCK2_ANDMA
ID   SCK2_ANDMA     STANDARD;      PRT;    59 AA.
AC   P45696;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Kaliotoxin 2 precursor (KTX 2).
GN   KTX2.
OS   Androctonus mauretanicus mauretanicus (Scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Androctonus.
OX   NCBI_TaxID=6860;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-59.
RC   TISSUE=Venom;
RX   MEDLINE=95105165; PubMed=7806508;
RA   Laraba-Djebari F., Legros C., Crest M., Ceard B., Romi R.,
RA   Mansuelle P., Jacquet G., van Rietschoten J., Gola M., Rochat H.;
RT   "The kaliotoxin family enlarged. Purification, characterization, and
RT   precursor nucleotide sequence of KTX2 from Androctonus australis
RT   venom.";
RL   J. Biol. Chem. 269:32835-32843(1994).
CC   -!- FUNCTION: INHIBITOR OF THE DENDROTOXIN SENSITIVE VOLTAGE-DEPENDENT
CC       K+ CHANNEL AND OF CALCIUM-ACTIVATED POTASSIUM CHANNELS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM
CC       CHANNEL INHIBITORS SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S74733; AAB33535.1; -.
DR   HSSP; P55896; 1SCO.
DR   InterPro; IPR001947; Scorpion_toxinS.
DR   Pfam; PF00451; toxin_2; 1.
DR   ProDom; PD003586; Scorpion_toxinS; 1.
DR   PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KW   Neurotoxin; Potassium channel inhibitor; Signal.
```

```
FT   SIGNAL        1     22
FT   CHAIN        23     59       KALIOTOXIN 2.
FT   DISULFID     29     49       BY SIMILARITY.
FT   DISULFID     35     54       BY SIMILARITY.
FT   DISULFID     39     56       BY SIMILARITY.
FT   SITE         47     54       INTERACTION WITH CA(2+)-ACTIVATED K(+)
FT                                CHANNELS (POTENTIAL).
SQ   SEQUENCE   59 AA;  6393 MW;  DD3225C5D270EB26 CRC64;

  Query Match             10.0%;  Score 44;  DB 1;  Length 59;
  Best Local Similarity   23.3%;  Pred. No. 1.9e+02;
  Matches   14;  Conservative    4;  Mismatches   20;  Indels   22;  Gaps    1;

Qy       5 VLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCK 64
           | |::| || :|  |      | |                          |  :  ||||
Db       3 VFSAVLIILFVCSMIIGINAVRIPVSC----------------------KHSGQCLKPCK 40


Search completed: May  8, 2003, 21:46:54
Job time : 12 secs
```

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:44:33 ; Search time 21 Seconds
(without alignments)
370.804 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 21269

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	59	13.3	55	2	E58892	H+-transporting tw
2	58	13.1	76	2	T43204	hypothetical prote
3	54.5	12.3	75	2	D81239	hypothetical prote
4	53	12.0	56	2	B75146	ssu ribosomal prot
5	53	12.0	56	2	H71185	probable ribosomal
6	53	12.0	64	2	AC2811	hypothetical prote
7	51	11.5	79	2	C82382	hypothetical prote
8	50	11.3	42	2	T07557	hypothetical prote
9	48.5	11.0	78	2	T11918	ribosomal protein
10	48	10.9	61	2	F86885	30S ribosomal prot
11	48	10.9	61	2	I59337	mammary transformi
12	48	10.9	69	2	T47058	hypothetical prote
13	48	10.9	74	1	GSFF7	salivary glue prot
14	48	10.9	79	2	T20201	hypothetical prote
15	47.5	10.7	77	2	F90702	hypothetical prote
16	47.5	10.7	77	2	A85553	hypothetical prote
17	47	10.6	69	2	T44065	ribosomal protein
18	47	10.6	73	4	S00695	hypothetical prote
19	45.5	10.3	66	2	PQ0360	spike glycoprotein
20	45.5	10.3	77	2	C82409	hypothetical prote
21	45	10.2	44	2	G37286	olfactory receptor
22	45	10.2	75	1	GSFF8	salivary glue prot
23	44.5	10.1	63	2	A99321	hypothetical prote
24	44.5	10.1	66	2	PQ0356	spike glycoprotein
25	44.5	10.1	66	2	PQ0354	spike glycoprotein
26	44.5	10.1	66	2	PQ0358	spike glycoprotein
27	44.5	10.1	66	2	PQ0348	spike glycoprotein
28	44.5	10.1	66	2	PQ0350	spike glycoprotein
29	44.5	10.1	77	2	G64784	ybcJ protein - Esc
30	44	10.0	59	1	A55807	kaliotoxin 2 precu
31	44	10.0	60	2	T07288	hydroxysteroid sul
32	44	10.0	63	2	A45041	trypsin inhibitor
33	44	10.0	67	2	AG0383	hypothetical prote
34	43.5	9.8	52	2	C82806	hypothetical prote
35	43.5	9.8	66	2	PQ0342	spike glycoprotein
36	43.5	9.8	73	2	AB1897	hypothetical prote
37	43.5	9.8	75	2	S22329	gene E protein - p
38	43.5	9.8	77	2	G96908	hypothetical prote
39	43	9.7	34	2	S42051	GTP-binding protei
40	43	9.7	48	2	S27128	histone H1 - sea u
41	43	9.7	49	2	S02007	protamine I - rabb
42	43	9.7	54	2	S28711	hypothetical prote
43	43	9.7	54	2	S34202	probable membrane
44	43	9.7	67	2	D82544	hypothetical prote
45	43	9.7	69	2	G95002	hypothetical prote

ALIGNMENTS

RESULT 1
E58892
H+-transporting two-sector ATPase (EC 3.6.3.14) protein 8 - coelacanth mitochondrion
C;Species: mitochondrion Latimeria chalumnae (coelacanth)
C;Date: 30-Oct-1998 #sequence_revision 30-Oct-1998 #text_change 03-Jun-2002
C;Accession: E58892
R;Zardoya, R.; Meyer, A.
Genetics 146, 995-1010, 1997
A;Title: The complete DNA sequence of the mitochondrial genome of a "living fossil,"
A;Reference number: A58892; MUID:97358858; PMID:9215903
A;Accession: E58892
A;Status: preliminary; nucleic acid sequence not shown; translation not shown; not c
A;Molecule type: DNA
A;Residues: 1-55 <ZAR>
A;Cross-references: GB:U82228; NID:g1916817; PIDN:AAC60322.1; PID:g1916822
A;Note: submitted to GenBank/EMBL/DDBJ December, 1996
C;Genetics:
A;Gene: ATPase8
A;Map position: FOR7944-8111
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: H+-transporting ATP synthase protein 8
C;Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion;

Query Match 13.3%; Score 59; DB 2; Length 55;
Best Local Similarity 34.3%; Pred. No. 14;
Matches 23; Conservative 7; Mismatches 15; Indels 22; Gaps 5;

```
Qy       7 SSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPS-PNSTNLKGHHVRLCKPCKL 65
           |  | |||  : || |     |:|      |:|  |  |: |:: |:          ||
Db       7 SPWLLILLFSWLIFLT---MLPSK------TQL--HTFPNMPSTQNM----------CKQ 45

Qy      66 EPEPRLW 72
           ||||  |
Db      46 EPEPWTW 52
```

RESULT 2
T43204
hypothetical protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T43204
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T43204
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-76 <YOS>
A;Cross-references: EMBL:D89274; NID:g1749755; PIDN:BAA13935.1; PID:g1749756

```
A;Experimental source: strain PR745

  Query Match             13.1%;  Score 58;  DB 2;  Length 76;
  Best Local Similarity   42.9%;  Pred. No. 23;
  Matches   18;  Conservative    2;  Mismatches   18;  Indels    4;  Gaps    2;

Qy      24 GKRRPAKAWSGRRTRL---CCHRVPSPNSTNLKGHHVRLCKP 62
           | || :   |  | |    |  | | | || || :  ||| |
Db       5 GSRRSSCIVSNARKRQSVHCLSRSPIP-STRLKENDSRLCSP 45


RESULT 3
D81239
hypothetical protein NMB0092 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: D81239
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: D81239
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <TET>
A;Cross-references: GB:AE002368; GB:AE002098; NID:g7225303; PIDN:AAF40554.1; PID:g722530
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0092

  Query Match             12.3%;  Score 54.5;  DB 2;  Length 75;
  Best Local Similarity   29.7%;  Pred. No. 55;
  Matches   22;  Conservative    8;  Mismatches   25;  Indels   19;  Gaps    4;

Qy       1 MRLLVLSSL----LCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCH-RVPSPNSTNLKGH 55
           :|  ||| |    || |  | | |            :|    |  | :| :||  |
Db       2 VRFFVLSFLTLINLCSLSACNSHF-----------TGNINPLGTHNKVANPNCANSANS 49

Qy      56 HVRLCKPCKLEPEP 69
           |:|  :| :   :|
Db      50 HIR--QPSRKNYDP 61


RESULT 4
B75146
ssu ribosomal protein s14p (rps14p) PAB7080 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: B75146
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: B75146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-56 <KAW>
A;Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49249.1; PID:g545775
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB7080
C;Superfamily: Escherichia coli ribosomal protein S14

  Query Match             12.0%;  Score 53;  DB 2;  Length 56;
  Best Local Similarity   29.8%;  Pred. No. 63;
  Matches   14;  Conservative    8;  Mismatches   21;  Indels    4;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRL 71
           ||:| |   | |  : | :        :: | : ||: |  |  |:|
Db       8 KRKPRKFGKGARRCIRCGQY----GPIIRIHGLMLCRHCFREVAPKL 50


RESULT 5
H71185
probable ribosomal protein S14 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: H71185
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophili
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: H71185
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-56 <KAW>
A;Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30879.1; PID:g3258196
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBa
C;Genetics:
A;Gene: PHS047
C;Superfamily: Escherichia coli ribosomal protein S14

  Query Match             12.0%;  Score 53;  DB 2;  Length 56;
  Best Local Similarity   29.8%;  Pred. No. 63;
  Matches   14;  Conservative    8;  Mismatches   21;  Indels    4;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRL 71
           ||:| |   | |  : | :        :: | : ||: |  |  |:|
Db       8 KRKPRKFGKGARRCIRCGQY----GPIIRIHGLMLCRHCFREVAPKL 50


RESULT 6
AC2811
hypothetical protein Atu1909 [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AC2811
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AC2811
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-64 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42905.1; PID:g17740360; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1909
A;Map position: circular chromosome

  Query Match             12.0%;  Score 53;  DB 2;  Length 64;
  Best Local Similarity   28.0%;  Pred. No. 71;
  Matches   14;  Conservative    6;  Mismatches   16;  Indels   14;  Gaps    1;

Qy      25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVV 74
           |:|||:              | | | : ||||  : | :|  | :
Db      17 KKRPARR-------------PVKKTALSGGYRRLCKGMQWENDPTEWSI 52


RESULT 7
C82382
hypothetical protein VCA1066 [imported] - Vibrio cholerae (strain N16961 serogroup O1
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: C82382
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
```

```
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82382
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <HEI>
A;Cross-references: GB:AE004432; GB:AE003853; NID:g9658509; PIDN:AAF96960.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA1066
A;Map position: 2

  Query Match            11.5%;  Score 51;  DB 2;  Length 79;
  Best Local Similarity  42.9%;  Pred. No. 1.4e+02;
  Matches    9;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      39 LCCHRVPSPNSTNLKGHHVRL 59
           | ||   :|:|  : | ||:|
Db      50 LTCHLPATPSSLGIAGSHVKL 70


RESULT 8
T07557
hypothetical protein 42f - Japanese black pine chloroplast
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000
C;Accession: T07557
R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome
A;Reference number: Z16030; MUID:95024047; PMID:7937893
A;Accession: T07557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-42 <WAK>
A;Cross-references: EMBL:D17510; NID:g529643; PIDN:BAA04433.1; PID:g1262718
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

  Query Match            11.3%;  Score 50;  DB 2;  Length 42;
  Best Local Similarity  37.5%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      63 CKLEPEPRLWVVPGAL 78
           |   |:| :|:: | |
Db       6 CSSAPDPEMWIIQGTL 21


RESULT 9
T11918
ribosomal protein S19 - Prototheca wickerhamii mitochondrion
C;Species: mitochondrion Prototheca wickerhamii
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
C;Accession: T11918
R;Wolff, G.; Plante, I.; Lang, B.F.; Kueck, U.; Burger, G.
J. Mol. Biol. 237, 75-86, 1994
A;Title: Complete sequence of the mitochondrial DNA of the chlorophyte alga Prototheca w
A;Reference number: Z17373; MUID:94180393; PMID:8133522
A;Accession: T11918
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-78 <WOL>
A;Cross-references: EMBL:U02970; NID:g467843; PID:g467850; PIDN:AAD12637.1
A;Experimental source: strain 263-11
C;Genetics:
A;Gene: rps19
A;Genome: mitochondrion
C;Superfamily: Escherichia coli ribosomal protein S19
C;Keywords: mitochondrion; protein biosynthesis; ribosome

  Query Match            11.0%;  Score 48.5;  DB 2;  Length 78;
  Best Local Similarity  33.3%;  Pred. No. 2.6e+02;
  Matches   13;  Conservative    7;  Mismatches   14;  Indels    5;  Gaps    2;

Qy      30 KAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPE 68
           | || ||: :   :|      :  |: |:  |||: ||
Db      23 KIWS-RRSAI----LPQFVGKTVSIHNGRIFIPCKISPE 56


RESULT 10
F86885
30S ribosomal protein S14 [imported] - Lactococcus lactis subsp. lactis (strain IL140
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 02-Aug-2002
C;Accession: F86885
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: F86885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <STO>
A;Cross-references: GB:AE005176; PID:g12725139; PIDN:AAK06184.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: rpsN
C;Superfamily: Escherichia coli ribosomal protein S14

  Query Match            10.9%;  Score 48;  DB 2;  Length 61;
  Best Local Similarity  33.9%;  Pred. No. 2.4e+02;
  Matches   19;  Conservative    5;  Mismatches   22;  Indels   10;  Gaps    3;

Qy      26 RRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 81
           :|||  :   ||  | |   |:|   |    :||: |  |         | || |
Db      11 QRPAKFSTQAYTR--CERCGRPHSVYRK---FKLCRICLRE-----LAYKGQLPGV 56


RESULT 11
I59337
mammary transforming protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I59337
R;Bera, T.K.; Guzman, R.C.; Miyamoto, S.; Panda, D.K.; Sasaki, M.; Hanyu, K.; Enami,
Proc. Natl. Acad. Sci. U.S.A. 91, 9789-9793, 1994
A;Title: Identification of a mammary transforming gene (MAT1) associated with mouse m
A;Reference number: I59337; MUID:95024046; PMID:7937892
A;Accession: I59337
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-61 <RES>
A;Cross-references: GB:L31958; NID:g473909; PIDN:AAC37665.1; PID:g473910
C;Genetics:
A;Gene: MAT1

  Query Match            10.9%;  Score 48;  DB 2;  Length 61;
  Best Local Similarity  28.3%;  Pred. No. 2.4e+02;
  Matches   13;  Conservative    7;  Mismatches   16;  Indels   10;  Gaps    2;

Qy      17 FSIFSTEGKRRPAKAWSGR------RTRLCCHRVPSPNSTNLKGHH 56
           | :||     || : | |       | :|    : |||  ::  :|
Db      13 FVVFSINLLSRPEREWEGMPQKGSGRAKL----LQSPNRKHISTNH 54


RESULT 12
T47058
hypothetical protein [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C;Accession: T47058
R;Buchrieser, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carn
```

```
submitted to the EMBL Data Library, October 1998
A;Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A;Reference number: Z24348
A;Accession: T47058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-69 <BUC>
A;Cross-references: EMBL:AL031866; PIDN:CAA21401.1
A;Experimental source: strain 6/69

  Query Match             10.9%;  Score 48;  DB 2;  Length 69;
  Best Local Similarity   25.5%;  Pred. No. 2.6e+02;
  Matches   14;  Conservative    5;  Mismatches   34;  Indels    2;  Gaps    1;

Qy      16 CFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPR 70
           |      : : ||    | :  | |   :|      | |  |  |  |:   | |
Db      15 CLEALIEQSRLRPTPTPSLQHCRFCGKAIPEKRRQTLPG--VTTCTDCQSILEKR 67


RESULT 13
GSFF7
salivary glue protein sgs-7 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 16-Jul-1999
C;Accession: A03330
R;Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.
J. Mol. Biol. 168, 765-789, 1983
A;Title: DNA sequences, gene regulation and modular protein evolution in the Drosophila
A;Reference number: A92904; MUID:83294545; PMID:6411930
A;Accession: A03330
A;Molecule type: DNA
A;Residues: 1-74 <GAR>
A;Cross-references: GB:X01918; NID:g8581; PIDN:CAA25993.1; PID:g603988
C;Comment: This protein is produced by third-instar larvae.
C;Genetics:
A;Gene: sgs-7
A;Cross-references: FlyBase:FBgn0003377
A;Map position: 3L (68C)
A;Introns: 10/1
C;Superfamily: salivary glue protein
C;Keywords: salivary gland

  Query Match             10.9%;  Score 48;  DB 1;  Length 74;
  Best Local Similarity   26.8%;  Pred. No. 2.8e+02;
  Matches   22;  Conservative   18;  Mismatches   26;  Indels   16;  Gaps    7;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEG--KRRP----AKAWSG--RRTRLCCHRVPSPNSTNL 52
           |:|: :: : ||||: ||  :  |  : :|     || :|   : :||   :    :
Db       1 MKLIAVTIIACILLIGFSDLALGGACECQPCGPGGKACTGCPEKPQLCQQLISDIRNLQQ 60

Qy      53 KGHHVRLCKPCKLEPEPRLWVV 74
           |   :| |  |  ||:   |::
Db      61 K---IRKC-VCG-EPQ---WMI 74


RESULT 14
T20201
hypothetical protein C54C8.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20201
R;Dobson, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19237
A;Accession: T20201
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-79 <WIL>
A;Cross-references: EMBL:Z83102; PIDN:CAB05463.1; GSPDB:GN00019; CESP:C54C8.9
A;Experimental source: clone C54C8
C;Genetics:
A;Gene: CESP:C54C8.9
A;Map position: 1

  Query Match             10.9%;  Score 48;  DB 2;  Length 79;
  Best Local Similarity   30.4%;  Pred. No. 2.9e+02;
  Matches   24;  Conservative   11;  Mismatches   26;  Indels   18;  Gaps    5;

Qy       1 MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHR-------VPSPNST 50
           |:||:| ||  |    | : : :    |    :|| ||:  ||          ||:  :
Db       1 MKLLILFSLFAIF---FGVIALDSP-IPEFYSTGRSTRVPSHRHRIPRLGKREVPNFQAD 56

Qy      51 NL--KGHHVR--LCKPCKL 65
           |:   |  ||  :  | ||
Db      57 NVPEAGGRVRRYVGPPLKL 75


RESULT 15
F90702
hypothetical protein ECs0590 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: F90702
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90702
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34013.1; PID:g13360048; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0590

  Query Match             10.7%;  Score 47.5;  DB 2;  Length 77;
  Best Local Similarity   36.6%;  Pred. No. 3.3e+02;
  Matches   15;  Conservative    5;  Mismatches   16;  Indels    5;  Gaps    2;

Qy      42 HRVPSPNSTNLKGH-HVRLCKPCKLEPEPRLWVVPGALPQV 81
           ||| :  : :|  | || ||   |||      |   ||  ::
Db       3 HRVSNMATFSLGKHPHVELCDLLKLEG----WSESGAQAKI 39


Search completed: May  8, 2003, 21:48:09
Job time : 22 secs
```

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 22:39:38 ; Search time 104 Seconds
(without alignments)
967.876 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 238982

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113853_7721/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500
-MINLEN=25 -MAXLEN=80 -USER=US09599087_@CGN_1_1_93_@runat_07052003_113853_7721
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	44	10.0	50	9	US-10-126-617-7	Sequence 7, Appli
c	2	44	10.0	50	9	US-10-126-617-8	Sequence 8, Appli
	3	43	9.7	69	9	US-10-159-428-1	Sequence 1, Appli
	4	43	9.7	78	9	US-10-204-200-6	Sequence 6, Appli
c	5	43	9.7	78	9	US-10-204-200-7	Sequence 7, Appli
	6	43	9.7	78	10	US-09-950-844-5	Sequence 5, Appli
c	7	43	9.7	78	10	US-09-950-844-6	Sequence 6, Appli
	8	43	9.7	78	10	US-09-835-699-5	Sequence 5, Appli
c	9	43	9.7	78	10	US-09-835-699-6	Sequence 6, Appli
c	10	42.5	9.6	59	9	US-10-006-593-89	Sequence 89, Appl
	11	42	9.5	50	9	US-10-137-788-6	Sequence 6, Appli
c	12	41.5	9.4	46	10	US-09-827-864-19	Sequence 19, Appl
c	13	41.5	9.4	54	10	US-09-294-093B-4383	Sequence 4383, Ap
	14	41.5	9.4	60	9	US-09-902-941-1979	Sequence 1979, Ap
	15	41.5	9.4	60	9	US-10-017-754-1979	Sequence 1979, Ap
	16	41	9.3	60	9	US-09-954-692-13	Sequence 13, Appl
	17	41	9.3	60	10	US-09-559-671A-13	Sequence 13, Appl
	18	41	9.3	80	9	US-09-440-829-17	Sequence 17, Appl
c	19	41	9.3	80	10	US-09-157-748-43	Sequence 43, Appl
c	20	40.5	9.2	75	10	US-09-864-761-28813	Sequence 28813, A
c	21	40	9.0	31	9	US-09-773-599-16	Sequence 16, Appl
c	22	40	9.0	31	10	US-09-801-274-1511	Sequence 1511, Ap
c	23	40	9.0	31	10	US-09-801-274-1512	Sequence 1512, Ap
	24	40	9.0	46	9	US-09-940-244-227	Sequence 227, App
	25	40	9.0	48	9	US-09-877-478-5348	Sequence 5348, Ap
c	26	40	9.0	59	9	US-10-025-380-976	Sequence 976, App
c	27	40	9.0	59	10	US-09-922-217-976	Sequence 976, App
c	28	40	9.0	59	10	US-09-833-263-976	Sequence 976, App
c	29	40	9.0	59	12	US-10-013-737-15	Sequence 15, Appl
	30	40	9.0	60	9	US-09-902-941-1980	Sequence 1980, Ap
	31	40	9.0	60	9	US-10-017-754-1980	Sequence 1980, Ap
	32	40	9.0	63	10	US-09-923-876-1639	Sequence 1639, Ap
c	33	40	9.0	64	9	US-10-072-036-7	Sequence 7, Appli
c	34	40	9.0	75	9	US-09-979-999-15	Sequence 15, Appl
	35	40	9.0	75	9	US-10-045-465-2	Sequence 2, Appli
	36	40	9.0	78	10	US-09-864-761-21687	Sequence 21687, A
	37	39.5	8.9	78	9	US-09-861-257-47	Sequence 47, Appl
	38	39	8.8	25	9	US-10-215-112-13778	Sequence 13778, A
	39	39	8.8	28	9	US-10-156-634A-13	Sequence 13, Appl
	40	39	8.8	36	9	US-09-345-373-5	Sequence 5, Appli
	41	39	8.8	37	10	US-09-263-959-815	Sequence 815, App
	42	39	8.8	40	10	US-09-263-959-458	Sequence 458, App
	43	39	8.8	64	9	US-10-085-906-168	Sequence 168, App
c	44	39	8.8	65	10	US-09-747-003-12	Sequence 12, Appl
	45	39	8.8	69	9	US-10-076-816-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-10-126-617-7/c
; Sequence 7, Application US/10126617
; Patent No. US20020168723A1
; GENERAL INFORMATION:
; APPLICANT: KAMBOJ, RAJENDER
; APPLICANT: ELLIOTT, CANDACE E.
; APPLICANT: NUTT, STEPHEN L.
; TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF THE EAA4 FAMILY
; FILE REFERENCE: 016777/0471
; CURRENT APPLICATION NUMBER: US/10/126,617
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 08/249,241
; PRIOR FILING DATE: 1994-05-25
; PRIOR APPLICATION NUMBER: 07/903,456
; PRIOR FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-617-7

Alignment Scores:
Pred. No.: 566 Length: 50
Score: 44.00 Matches: 7
Percent Similarity: 60.00% Conservative: 2

Best Local Similarity: 46.67% Mismatches: 6
Query Match: 9.95% Indels: 0
DB: 9 Gaps: 0

US-09-599-087B-5 (1-81) x US-10-126-617-7 (1-50)

```
Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||     :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2
```

RESULT 2
US-10-126-617-8/c
; Sequence 8, Application US/10126617
; Patent No. US20020168723A1
; GENERAL INFORMATION:
; APPLICANT: KAMBOJ, RAJENDER
; APPLICANT: ELLIOTT, CANDACE E.
; APPLICANT: NUTT, STEPHEN L.
; TITLE OF INVENTION: KAINATE-BINDING HUMAN CNS RECEPTORS OF THE EAA4 FAMILY
; FILE REFERENCE: 016777/0471
; CURRENT APPLICATION NUMBER: US/10/126,617
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 08/249,241
; PRIOR FILING DATE: 1994-05-25
; PRIOR APPLICATION NUMBER: 07/903,456
; PRIOR FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-617-8

Alignment Scores:
Pred. No.: 566 Length: 50
Score: 44.00 Matches: 7
Percent Similarity: 60.00% Conservative: 2
Best Local Similarity: 46.67% Mismatches: 6
Query Match: 9.95% Indels: 0
DB: 9 Gaps: 0

US-09-599-087B-5 (1-81) x US-10-126-617-8 (1-50)

```
Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||     :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2
```

RESULT 3
US-10-159-428-1
; Sequence 1, Application US/10159428
; Publication No. US20030082579A1
; GENERAL INFORMATION:
; APPLICANT: Felgner, Philip L
; APPLICANT: Doolan, Denise L
; TITLE OF INVENTION: PROTEIN ARRAYS AND METHODS AND SYSTEMS
; TITLE OF INVENTION: FOR PRODUCING THE SAME
; FILE REFERENCE: GTSYS.008A
; CURRENT APPLICATION NUMBER: US/10/159,428
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/294,739
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-159-428-1

Alignment Scores:
Pred. No.: 1.13e+03 Length: 69
Score: 43.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 9.73% Indels: 0
DB: 9 Gaps: 0

US-09-599-087B-5 (1-81) x US-10-159-428-1 (1-69)

```
Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||||   ::::::|||
Db      22 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 57
```

RESULT 4
US-10-204-200-6
; Sequence 6, Application US/10204200
; Publication No. US20030036108A1
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Kathleen A.
; APPLICANT: Benincasa, Diana
; APPLICANT: Cascieri, Margaret A.
; APPLICANT: Mitnaul, Lyndon J.
; APPLICANT: Shiao, Lin-Lin
; APPLICANT: Tota, Michael A.
; TITLE OF INVENTION: CELL FUSION ASSAYS USING FLUORESCENCE
; TITLE OF INVENTION: RESONANCE ENERGY TRANSFER
; FILE REFERENCE: 20618P
; CURRENT APPLICATION NUMBER: US/10/204,200
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: PCT/US01/04677
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/183,309
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Human
US-10-204-200-6

Alignment Scores:
Pred. No.: 1.32e+03 Length: 78
Score: 43.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 9.73% Indels: 0
DB: 9 Gaps: 0

US-09-599-087B-5 (1-81) x US-10-204-200-6 (1-78)

```
Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||||   ::::::|||
Db      29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64
```

RESULT 5
US-10-204-200-7/c
; Sequence 7, Application US/10204200
; Publication No. US20030036108A1
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Kathleen A.
; APPLICANT: Benincasa, Diana
; APPLICANT: Cascieri, Margaret A.
; APPLICANT: Mitnaul, Lyndon J.
; APPLICANT: Shiao, Lin-Lin
; APPLICANT: Tota, Michael A.
; TITLE OF INVENTION: CELL FUSION ASSAYS USING FLUORESCENCE
; TITLE OF INVENTION: RESONANCE ENERGY TRANSFER
; FILE REFERENCE: 20618P
; CURRENT APPLICATION NUMBER: US/10/204,200
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: PCT/US01/04677
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/183,309
; PRIOR FILING DATE: 2000-02-17

```
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Human
US-10-204-200-7

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-204-200-7 (1-78)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::||||||||||||||   :::::::|||
Db     54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19

RESULT 6
US-09-950-844-5
; Sequence 5, Application US/09950844
; Patent No. US20020045594A1
; GENERAL INFORMATION:
;  APPLICANT: Volkin, David B.
;  APPLICANT:  Evans, Robert K.
;  APPLICANT:  Ulmer, Jeffrey B.
;  APPLICANT:  Caulfield, Michael J.
;  TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULATIONS
;  FILE REFERENCE: 19907YIACA
;  CURRENT APPLICATION NUMBER: US/09/950,844
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: 09/112,655
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: 09/023,834
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 60/038,194
;  PRIOR FILING DATE: 1997-02-14
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-950-844-5

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-950-844-5 (1-78)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::||||||||||||||   :::::::|||
Db     29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 7
US-09-950-844-6/c
; Sequence 6, Application US/09950844
; Patent No. US20020045594A1
; GENERAL INFORMATION:
;  APPLICANT: Volkin, David B.
;  APPLICANT:  Evans, Robert K.
;  APPLICANT:  Ulmer, Jeffrey B.
;  APPLICANT:  Caulfield, Michael J.
;  TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULATIONS
;  FILE REFERENCE: 19907YIACA
;  CURRENT APPLICATION NUMBER: US/09/950,844
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: 09/112,655
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: 09/023,834
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 60/038,194
;  PRIOR FILING DATE: 1997-02-14
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-950-844-6

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-950-844-6 (1-78)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||   |||:::||||||||||||||   :::::::|||
Db     54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19

RESULT 8
US-09-835-699-5
; Sequence 5, Application US/09835699
; Patent No. US20020147167A1
;  GENERAL INFORMATION:
;       APPLICANT: Armstrong, Marcy E.
;                  Keys, Robert D.
;                  Lewis, John A.
;                  Liu, Margaret A.
;                  McClements, William L.
;       TITLE OF INVENTION: A POLYNUCLEOTIDE HERPES VIRUS
;                           VACCINE
;       NUMBER OF SEQUENCES: 15
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Merck & Co., Inc.
;            STREET: P.O. Box 2000, 126 E. Lincoln Ave.
;            CITY: Rahway
;            STATE: NJ
;            COUNTRY: USA
;            ZIP: 07065-0907
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: Windows
;            SOFTWARE: FastSEQ for Windows Version 2.0b
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/835,699
;            FILING DATE: 17-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/US97/17262
;            FILING DATE: 26-SEP-1997
;            APPLICATION NUMBER: US 08/720,758
;            FILING DATE: 01-OCT-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hand, J. Mark
;            REGISTRATION NUMBER: 36,545
;            REFERENCE/DOCKET NUMBER: 19258CC
```

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 732-594-3905
;             TELEFAX: 732-594-4720
;             TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 5:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 78 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         MOLECULE TYPE: Other
;         FEATURE:
;             NAME/KEY:
;             LOCATION:  1...78
;             OTHER INFORMATION:  nucleic acid
;         SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-835-699-5

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-835-699-5 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||   |||:::||||||||||||   ::::::|||
Db      29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 9
US-09-835-699-6/c
; Sequence 6, Application US/09835699
; Patent No. US20020147167A1
;   GENERAL INFORMATION:
;         APPLICANT: Armstrong, Marcy E.
;                    Keys, Robert D.
;                    Lewis, John A.
;                    Liu, Margaret A.
;                    McClements, William L.
;         TITLE OF INVENTION: A POLYNUCLEOTIDE HERPES VIRUS
;                             VACCINE
;         NUMBER OF SEQUENCES: 15
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Merck & Co., Inc.
;              STREET: P.O. Box 2000, 126 E. Lincoln Ave.
;              CITY: Rahway
;              STATE: NJ
;              COUNTRY: USA
;              ZIP: 07065-0907
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Diskette
;              COMPUTER: IBM Compatible
;              OPERATING SYSTEM: Windows
;              SOFTWARE: FastSEQ for Windows Version 2.0b
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/835,699
;              FILING DATE: 17-Aug-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: PCT/US97/17262
;              FILING DATE: 26-SEP-1997
;              APPLICATION NUMBER: US 08/720,758
;              FILING DATE: 01-OCT-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Hand, J. Mark
;              REGISTRATION NUMBER: 36,545
;              REFERENCE/DOCKET NUMBER: 19258CC
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 732-594-3905
;              TELEFAX: 732-594-4720
;              TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 78 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Other
;         FEATURE:
;              NAME/KEY:
;              LOCATION:  1...78
;              OTHER INFORMATION:  nucleic acid
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-835-699-6

Alignment Scores:
Pred. No.:              1.32e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x US-09-835-699-6 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||   |||:::||||||||||||   ::::::|||
Db      54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19

RESULT 10
US-10-006-593-89/c
; Sequence 89, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
;  FILE REFERENCE: 1087-2
;  CURRENT APPLICATION NUMBER: US/10/006,593
;  CURRENT FILING DATE:  2001-12-05
;  PRIOR APPLICATION NUMBER: US 60/251,448
;  PRIOR FILING DATE: 2000-12-05
;  PRIOR APPLICATION NUMBER: US 60/288,889
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/294,068
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-006-593-89

Alignment Scores:
Pred. No.:              1.07e+03        Length:         59
Score:                  42.50           Matches:        9
Percent Similarity:     57.89%          Conservative:   2
Best Local Similarity:  47.37%          Mismatches:     7
Query Match:            9.62%           Indels:         1
DB:                     9               Gaps:           1

US-09-599-087B-5 (1-81) x US-10-006-593-89 (1-59)

Qy      15 LeuCysPheSerIlePheSerThrGluGlyLysArgArgPro---AlaLysAlaTrp 32
           |||   ||||||:::|||   |||      |||   :::|||   |||      |||
Db      59 CTGCTATTTTCATTATTCAAAACCCGATCGAAGGCCCAACCCTGCGCCAGTGGCTGG 3

RESULT 11
US-10-137-788-6
```

```
; Sequence 6, Application US/10137788
; Publication No. US20030039636A1
; GENERAL INFORMATION:
;  APPLICANT: Leboulch, Philippe
;  APPLICANT:  Westerman, Karen
;  TITLE OF INVENTION: NOVEL SELF-INACTIVATING (SIN) LENTIVIRAL VECTORS
;  FILE REFERENCE: IOI-023
;  CURRENT APPLICATION NUMBER: US/10/137,788
;  CURRENT FILING DATE:  2002-05-01
;  PRIOR APPLICATION NUMBER: US 60/288,042
;  PRIOR FILING DATE: 2001-05-01
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
US-10-137-788-6

Alignment Scores:
Pred. No.:              1.01e+03        Length:         50
Score:                  42.00           Matches:        7
Percent Similarity:     73.33%          Conservative:   4
Best Local Similarity:  46.67%          Mismatches:     4
Query Match:            9.50%           Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x US-10-137-788-6 (1-50)

Qy      4 LeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSer 18
          |||   ||||||:::   :::|||:::|||:::   |||   |||
Db      4 CTTGCCTTGAGTGCTTCAATGTGTGTTGGTTTTTTGTGTGTCG 48

RESULT 12
US-09-827-864-19/c
; Sequence 19, Application US/09827864
; Patent No. US20020009458A1
;   GENERAL INFORMATION:
;        APPLICANT: COLAU, DIDIER
;                   ROOS, JOEL
;        TITLE OF INVENTION: RECOMBINANT DNA SEQUENCES, MOLECULES,
;                            VECTORS AND VACCINES FOR FELINE
;                            CALICIVIRUS DISEASE AND METHODS FOR
;                            PRODUCING AND USING SAME.
;        NUMBER OF SEQUENCES: 24
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MCDERMOTT, WILL & EMERY
;             STREET: 1850 K STREET, N.W., SUITE 500
;             CITY: WASHINGTON
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20006
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/827,864
;             FILING DATE: 06-Apr-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: GADIANO, WILHLEM F
;             REGISTRATION NUMBER: 37,136
;             REFERENCE/DOCKET NUMBER: 37712-213
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 778-8373
;             TELEFAX: (202) 778-8335
;   INFORMATION FOR SEQ ID NO: 19:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 46 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: nucleic acid
;        SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-827-864-19

Alignment Scores:
Pred. No.:              1.05e+03        Length:         46
Score:                  41.50           Matches:        8
Percent Similarity:     64.29%          Conservative:   1
Best Local Similarity:  57.14%          Mismatches:     4
Query Match:            9.39%           Indels:         1
DB:                     10              Gaps:           1

US-09-599-087B-5 (1-81) x US-09-827-864-19 (1-46)

Qy      60 CysLysProCysLysLeuGluProGluProArgLeuTrpVal 73
           |||   ||||||||||||||||      |||      |||:::
Db      42 TGTCACCCTTGCAAGCTGGAATGACGTCCTTCT---TGGTTA 4

RESULT 13
US-09-294-093B-4383/c
; Sequence 4383, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:      PERL Program
; SEQ ID NO 4383
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700353949H1
;   NAME/KEY: unsure
;   LOCATION: 34
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4383

Alignment Scores:
Pred. No.:              1.28e+03        Length:         54
Score:                  41.50           Matches:        8
Percent Similarity:     69.23%          Conservative:   1
Best Local Similarity:  61.54%          Mismatches:     1
Query Match:            9.39%           Indels:         3
DB:                     10              Gaps:           1

US-09-599-087B-5 (1-81) x US-09-294-093B-4383 (1-54)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSer 46
           |||||||||         |||||||||   :::|||||||
Db      52 GGACGGAGG---------TGCTGCCATNTATTACCAAGT 23

RESULT 14
US-09-902-941-1979
; Sequence 1979, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
```

; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1979
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1979

Alignment Scores:

Pred. No.:	1.46e+03	Length:	60
Score:	41.50	Matches:	9
Percent Similarity:	61.11%	Conservative:	2
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	9.39%	Indels:	1
DB:	9	Gaps:	1

US-09-599-087B-5 (1-81) x US-09-902-941-1979 (1-60)

```
Qy      34 GlyArgArgThrArgLeu---CysCysHisArgValProSerProAsnSerThr 50
           |||   :::   :::|||   ||||||         |||||||||   ||||||
Db       1 GGGGCTCAGGCGAAGCTGGGGTGCTGTTGGGGGTATCCGAGTCCCAGAAGCACC 54
```

RESULT 15
US-10-017-754-1979
; Sequence 1979, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1979
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1979

Alignment Scores:

Pred. No.:	1.46e+03	Length:	60
Score:	41.50	Matches:	9
Percent Similarity:	61.11%	Conservative:	2
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	9.39%	Indels:	1
DB:	9	Gaps:	1

US-09-599-087B-5 (1-81) x US-10-017-754-1979 (1-60)

```
Qy      34 GlyArgArgThrArgLeu---CysCysHisArgValProSerProAsnSerThr 50
           |||   :::   :::|||   ||||||         |||||||||   ||||||
Db       1 GGGGCTCAGGCGAAGCTGGGGTGCTGTTGGGGGTATCCGAGTCCCAGAAGCACC 54
```

Search completed: May 8, 2003, 23:29:51
Job time : 104 secs

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 22:37:43 ; Search time 62 Seconds
(without alignments)
400.658 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 322676

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113853_7693/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_36_@runat_07052003_113853_7693 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		53	12.0	79	1	US-08-170-290A-19	Sequence 19, Appl
2		46.5	10.5	48	2	US-08-350-260A-289	Sequence 289, App
3	c	46	10.4	72	1	US-08-441-430-22	Sequence 22, Appl
4		45.5	10.3	69	1	US-08-554-612C-37	Sequence 37, Appl
5		45	10.2	39	2	US-08-350-260A-233	Sequence 233, App
6		44	10.0	42	1	US-08-642-255-23	Sequence 23, Appl
7	c	44	10.0	50	1	US-07-903-456-7	Sequence 7, Appli
8	c	44	10.0	50	1	US-07-903-456-8	Sequence 8, Appli
9	c	44	10.0	50	4	US-08-249-241-7	Sequence 7, Appli
10	c	44	10.0	50	4	US-08-249-241-8	Sequence 8, Appli
11		44	10.0	51	2	US-08-350-260A-287	Sequence 287, App
12		44	10.0	51	2	US-08-350-260A-290	Sequence 290, App
13		43	9.7	72	6	5466668-7	Patent No. 5466668
14		43	9.7	78	1	US-08-338-992B-5	Sequence 5, Appli
15	c	43	9.7	78	1	US-08-338-992B-6	Sequence 6, Appli
16		43	9.7	78	4	US-09-010-733-5	Sequence 5, Appli
17	c	43	9.7	78	4	US-09-010-733-6	Sequence 6, Appli
18		43	9.7	78	5	PCT-US95-09057-5	Sequence 5, Appli
19	c	43	9.7	78	5	PCT-US95-09057-6	Sequence 6, Appli
20	c	43	9.7	79	1	US-08-571-983-2	Sequence 2, Appli
21	c	43	9.7	79	5	PCT-US94-08111-2	Sequence 2, Appli
22	c	42	9.5	30	1	US-08-338-471D-4	Sequence 4, Appli
23		42	9.5	69	1	US-07-977-284A-207	Sequence 207, App
24		42	9.5	69	2	US-08-256-426B-207	Sequence 207, App
25	c	41.5	9.4	46	4	US-08-872-056-19	Sequence 19, Appl
26		41.5	9.4	60	2	US-08-465-380-112	Sequence 112, App
27		41.5	9.4	60	2	US-08-480-478-27	Sequence 27, Appl
28		41.5	9.4	60	2	US-08-486-397-112	Sequence 112, App
29		41.5	9.4	60	2	US-08-486-399-112	Sequence 112, App
30		41.5	9.4	60	2	US-08-461-965-112	Sequence 112, App
31		41.5	9.4	60	2	US-08-326-110A-27	Sequence 27, Appl
32		41.5	9.4	60	2	US-08-634-641-112	Sequence 112, App
33		41.5	9.4	60	3	US-09-249-471-112	Sequence 112, App
34		41.5	9.4	60	3	US-09-249-472-112	Sequence 112, App
35		41.5	9.4	60	3	US-09-249-451-112	Sequence 112, App
36		41.5	9.4	60	3	US-08-809-455-112	Sequence 112, App
37		41.5	9.4	60	3	US-09-249-461-112	Sequence 112, App
38		41.5	9.4	60	3	US-09-249-448-112	Sequence 112, App
39		41	9.3	30	1	US-07-969-931-27	Sequence 27, Appl
40		41	9.3	30	1	US-07-855-417A-27	Sequence 27, Appl
41		41	9.3	60	4	US-09-339-913B-13	Sequence 13, Appl
42		41	9.3	60	4	US-09-339-904A-13	Sequence 13, Appl
43		41	9.3	60	4	US-08-769-062B-13	Sequence 13, Appl
44		41	9.3	60	4	US-09-344-002B-13	Sequence 13, Appl
45		41	9.3	60	4	US-09-559-565C-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-08-170-290A-19
; Sequence 19, Application US/08170290A
; Patent No. 5702931
;  GENERAL INFORMATION:
;    APPLICANT:  Andrews, William H.
;    APPLICANT:  Morser, Michael J.
;    APPLICANT:  Zielander, Laura R.
;    TITLE OF INVENTION:  No. 5702931el Mutagenesis Methods and
;    TITLE OF INVENTION:  Compositions
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James M. Heslin
;      STREET:  379 Lytton Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/170,290A
;      FILING DATE:  28-DEC-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/05573
;      FILING DATE:  01-JUL-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/724,237
;      FILING DATE:  01-JUL-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Heslin, James M.
;      REGISTRATION NUMBER:  29,541
;      REFERENCE/DOCKET NUMBER:  11972-58-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  79 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-170-290A-19

Alignment Scores:
Pred. No.:              114          Length:         79
Score:                  53.00        Matches:        9
Percent Similarity:     55.00%       Conservative:   2
Best Local Similarity:  45.00%       Mismatches:     9
Query Match:            11.99%       Indels:         0
DB:                     1            Gaps:           0

US-09-599-087B-5 (1-81) x US-08-170-290A-19 (1-79)

Qy      29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      |||:::|||   |||         |||   ||||||   |||   |||:::
Db      15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCTTGCCCGAGT 74

RESULT 2
US-08-350-260A-289
; Sequence 289, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  289:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-289

Alignment Scores:
Pred. No.:              320          Length:         48
Score:                  46.50        Matches:        9
Percent Similarity:     66.67%       Conservative:   3
Best Local Similarity:  50.00%       Mismatches:     3
Query Match:            10.52%       Indels:         3
DB:                     2            Gaps:           1

US-09-599-087B-5 (1-81) x US-08-350-260A-289 (1-48)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
           ||||||   ||||||   :::   ||||||         |||:::|||:::|||
Db       3 TGGTCCCTCCGCCGAATACAGTGCTGCTGT--------CCCACGCCTGACAGT 47

RESULT 3
US-08-441-430-22/c
; Sequence 22, Application US/08441430
; Patent No. 5681942
;  GENERAL INFORMATION:
;    APPLICANT:  Buchwald, Manuel
;    APPLICANT:  Strathdee, Craig A.
;    APPLICANT:  Wevrick, Rachel
;    APPLICANT:  Mathew, Christopher George Porter
;    TITLE OF INVENTION:  Fanconi Anemia Type C Gene
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Polley, Esq.
;      ADDRESSEE:  Klarquist, Sparkman, Campbell, Leigh &
;      ADDRESSEE:  Whinston, LLP
;      STREET:  121 S.W. Salmon, Suite 1600
;      CITY:  Portland
;      STATE:  Oregon
;      COUNTRY:  U.S.A.
;      ZIP:  97204
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk, 3+-inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect 5.1/ASCII Text File
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,430
;      FILING DATE:  May 15, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. 07/876,285
;      FILING DATE:  April 29, 1992
;      APPLICATION NUMBER:  U.S. 07/918,313
;      FILING DATE:  July 21, 1992
;      APPLICATION NUMBER:  U.S. 08/003,963
;      FILING DATE:  January 15, 1993
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Richard J. Polley, Esq.
;      REGISTRATION NUMBER:  28,107
;      REFERENCE/DOCKET NUMBER:  3812-42824
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (503) 226-7391
;      TELEFAX:  (503) 228-9446
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  72 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double stranded
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-441-430-22

Alignment Scores:
Pred. No.:              642          Length:         72
Score:                  46.00        Matches:        7
Percent Similarity:     76.92%       Conservative:   3
Best Local Similarity:  53.85%       Mismatches:     3
Query Match:            10.41%       Indels:         0
DB:                     1            Gaps:           0

US-09-599-087B-5 (1-81) x US-08-441-430-22 (1-72)

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
           |||||||||:::   :::|||  |||   |||:::|||
Db      71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 4
US-08-554-612C-37
; Sequence 37, Application US/08554612C
; Patent No. 5747660
;  GENERAL INFORMATION:
;    APPLICANT:  Orlicky, David
;    TITLE OF INVENTION:  PROSTAGLANDIN F2` RECEPTOR REGULATORY
;    TITLE OF INVENTION:  PROTEIN AND THERAPEUTIC USES
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  2200 Sand Hill Road, Suite 100
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/554,612C
;      FILING DATE:  No. 5747660ember 6, 1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  06519/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 322-5070
;      TELEFAX:  (415) 854-0875
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-554-612C-37

Alignment Scores:
Pred. No.:              691          Length:         69
Score:                  45.50        Matches:        10
Percent Similarity:     54.55%       Conservative:   2
Best Local Similarity:  45.45%       Mismatches:     7
Query Match:            10.29%       Indels:         3
DB:                     1            Gaps:           1

US-09-599-087B-5 (1-81) x US-08-554-612C-37 (1-69)

Qy      25 LysArgArgProAla---------LysAlaTrpSerGlyArgArgThrArgLeuCysCys 41
           :::|||||||   |||         :::|||||||      |||||||         |||||||
Db       3 AGGCGGAGAGTCGCTCCCGCCGGCCGAGCATGGGGCGCCCGGCGCCGAGGCCGCTGCTGC 62

Qy      42 HisArg 43
           |||
Db      63 TGGCGC 68

RESULT 5
US-08-350-260A-233
; Sequence 233, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  233:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-233

Alignment Scores:
Pred. No.:              357          Length:        39
Score:                  45.00        Matches:       8
Percent Similarity:     66.67%       Conservative:  0
Best Local Similarity:  66.67%       Mismatches:    4
Query Match:            10.18%       Indels:        0
DB:                     2            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-350-260A-233 (1-39)

Qy     32 TrpSerGlyArgArgThrArgLeuCysCysHisArg 43
          ||||||  ||||||        ||||||||||||
Db      3 TGGTCCCTCCGCCGAATACCAGGCTGCTGTCATAGG 38

RESULT 6
US-08-642-255-23
; Sequence 23, Application US/08642255
; Patent No. 5773249
;  GENERAL INFORMATION:
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic"
US-08-642-255-23

Alignment Scores:
Pred. No.:              516          Length:        42
Score:                  44.00        Matches:       7
Percent Similarity:     75.00%       Conservative:  2
Best Local Similarity:  58.33%       Mismatches:    3
Query Match:            9.95%        Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-642-255-23 (1-42)

Qy     27 ArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
          ||||||   ::::::|||      ||||||||||||
Db      4 AGGCCCAGGAGGTCCTGGCGGTCCAGGCGGACCAGG 39

RESULT 7
US-07-903-456-7/c
; Sequence 7, Application US/07903456
; Patent No. 5574144
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,456
;      FILING DATE:  19920624
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/183/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-07-903-456-7

Alignment Scores:
Pred. No.:              657          Length:        50
Score:                  44.00        Matches:       7
Percent Similarity:     60.00%       Conservative:  2
Best Local Similarity:  46.67%       Mismatches:    6
Query Match:            9.95%        Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-07-903-456-7 (1-50)

Qy     34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
          |||      :::      ||||||  |||  |||:::||||||
Db     46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 8
US-07-903-456-8/c
; Sequence 8, Application US/07903456
```

```
; Patent No. 5574144
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,456
;      FILING DATE:  19920624
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/183/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-07-903-456-8

Alignment Scores:
Pred. No.:              657          Length:        50
Score:                  44.00        Matches:       7
Percent Similarity:     60.00%       Conservative:  2
Best Local Similarity:  46.67%       Mismatches:    6
Query Match:            9.95%        Indels:        0
DB:                     1            Gaps:          0

US-09-599-087B-5 (1-81) x US-07-903-456-8 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 9
US-08-249-241-7/c
; Sequence 7, Application US/08249241
; Patent No. 6376200
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K. Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,241
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,456
;      FILING DATE:  24-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/220/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-249-241-7

Alignment Scores:
Pred. No.:              657          Length:        50
Score:                  44.00        Matches:       7
Percent Similarity:     60.00%       Conservative:  2
Best Local Similarity:  46.67%       Mismatches:    6
Query Match:            9.95%        Indels:        0
DB:                     4            Gaps:          0

US-09-599-087B-5 (1-81) x US-08-249-241-7 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 10
US-08-249-241-8/c
; Sequence 8, Application US/08249241
; Patent No. 6376200
;  GENERAL INFORMATION:
;    APPLICANT:  KAMBOJ, Rajender
;    APPLICANT:  ELLIOTT, Candace
;    APPLICANT:  NUTT, Stephen
;    TITLE OF INVENTION:  KAINATE-BINDING HUMAN CNS RECEPTORS OF
;    TITLE OF INVENTION:  THE EAA4 FAMILY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K. Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,241
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,456
;      FILING DATE:  24-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/220/ALLE
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-249-241-8

Alignment Scores:
Pred. No.:                657             Length:          50
Score:                    44.00           Matches:         7
Percent Similarity:       60.00%          Conservative:    2
Best Local Similarity:    46.67%          Mismatches:      6
Query Match:              9.95%           Indels:          0
DB:                       4               Gaps:            0

US-09-599-087B-5 (1-81) x US-08-249-241-8 (1-50)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||     :::     ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGTTGCATGAGAGCTCCAACTCCAAAC 2

RESULT 11
US-08-350-260A-287
; Sequence 287, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  287:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-287

Alignment Scores:
Pred. No.:                676             Length:          51
Score:                    44.00           Matches:         9
Percent Similarity:       62.50%          Conservative:    1
Best Local Similarity:    56.25%          Mismatches:      4
Query Match:              9.95%           Indels:          2
DB:                       2               Gaps:            1

US-09-599-087B-5 (1-81) x US-08-350-260A-287 (1-51)

Qy      32 TrpSerGlyArgArgThrArgLeuCys------CysHisArgValPro 45
           ||||||   ||||||   |||   |||      ||||||   :::|||
Db       3 TGGTCCCTCCGCCGAATACGAAGTTGCTCCCACTGCCATGGTCTGCCC 50

RESULT 12
US-08-350-260A-290
; Sequence 290, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  PCT/GB91/01134
;       FILING DATE:  10-JUL-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/GB92/00883
;       FILING DATE:  15-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/GB93/00605
;       FILING DATE:  24-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/150,002
;       FILING DATE:  31-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/307,619
;       FILING DATE:  16-SEP-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clough, David W
;       REGISTRATION NUMBER:  36,107
;       REFERENCE/DOCKET NUMBER:  28111/32372
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-474-6300
;   INFORMATION FOR SEQ ID NO:  290:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  51 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-350-260A-290

Alignment Scores:
Pred. No.:                676             Length:          51
Score:                    44.00           Matches:         9
Percent Similarity:       61.11%          Conservative:    2
Best Local Similarity:    50.00%          Mismatches:      5
Query Match:              9.95%           Indels:          2
DB:                       2               Gaps:            1

US-09-599-087B-5 (1-81) x US-08-350-260A-290 (1-51)

Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
           ||||||   ||||||         ||||||         |||:::|||:::|||
Db       3 TGGTCCCTCCGCCGAATACMNNAGTGCTGC------TGTCCCACGCCTGACAGT 50

RESULT 13
5466668-7
;Patent No. 5466668
;     APPLICANT: GLASER, CHARLES B.;MORSER, MICHAEL J.;LIGHT,
;DAVID R.
;     TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
;     NUMBER OF SEQUENCES: 57
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/155,346
;       FILING DATE: 22-NOV-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 568,456
;       FILING DATE: 15-AUG-1990
;       APPLICATION NUMBER: 506,325
;       FILING DATE: 09-APR-1990
;       APPLICATION NUMBER: 406,941
;       FILING DATE: 13-SEP-1989
;       APPLICATION NUMBER: 345,374
;       FILING DATE: 28-APR-1989
;SEQ ID NO:7:
;       LENGTH: 72
5466668-7

Alignment Scores:
Pred. No.:                1.42e+03        Length:          72
Score:                    43.00           Matches:         7
Percent Similarity:       83.33%          Conservative:    3
Best Local Similarity:    58.33%          Mismatches:      2
Query Match:              9.73%           Indels:          0
DB:                       6               Gaps:            0

US-09-599-087B-5 (1-81) x 5466668-7 (1-72)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||||   ::::::|||
Db      22 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 57

RESULT 14
US-08-338-992B-5
; Sequence 5, Application US/08338992B
; Patent No. 5736524
;  GENERAL INFORMATION:
;    APPLICANT:  CONTENT, JEAN
;    APPLICANT:  HUYGEN, KRIS
;    APPLICANT:  LIU, MARGARET A.
;    APPLICANT:  MONTGOMERY, DONNA
;    APPLICANT:  ULMER, JEFFREY
;    TITLE OF INVENTION:  A POLYNUCLEOTIDE TUBERCULOSIS VACCINE
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MICHAEL D. YABLONSKY - MERCK & CO., INC.
;      STREET:  126 E. LINCOLN AVE., - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/338,992B
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YABLONSKY, MICHAEL D.
;      REGISTRATION NUMBER:  40,407
;      REFERENCE/DOCKET NUMBER:  19342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-4678
;      TELEFAX:  908-594-4720
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-338-992B-5

Alignment Scores:
Pred. No.:                1.59e+03        Length:          78
Score:                    43.00           Matches:         7
Percent Similarity:       83.33%          Conservative:    3
Best Local Similarity:    58.33%          Mismatches:      2
Query Match:              9.73%           Indels:          0
DB:                       1               Gaps:            0

US-09-599-087B-5 (1-81) x US-08-338-992B-5 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
             |||   |||:::|||||||||||||   ::::::|||
Db      29 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 64

RESULT 15
US-08-338-992B-6/c
; Sequence 6, Application US/08338992B
; Patent No. 5736524
;  GENERAL INFORMATION:
;    APPLICANT:  CONTENT, JEAN
;    APPLICANT:  HUYGEN, KRIS
;    APPLICANT:  LIU, MARGARET A.
```

```
;    APPLICANT:  MONTGOMERY, DONNA
;    APPLICANT:  ULMER, JEFFREY
;    TITLE OF INVENTION:  A POLYNUCLEOTIDE TUBERCULOSIS VACCINE
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MICHAEL D. YABLONSKY - MERCK & CO., INC.
;      STREET:  126 E. LINCOLN AVE., - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/338,992B
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YABLONSKY, MICHAEL D.
;      REGISTRATION NUMBER:  40,407
;      REFERENCE/DOCKET NUMBER:  19342
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-4678
;      TELEFAX:  908-594-4720
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-338-992B-6

Alignment Scores:
Pred. No.:              1.59e+03        Length:         78
Score:                  43.00           Matches:        7
Percent Similarity:     83.33%          Conservative:   3
Best Local Similarity:  58.33%          Mismatches:     2
Query Match:            9.73%           Indels:         0
DB:                     1               Gaps:           0

US-09-599-087B-5 (1-81) x US-08-338-992B-6 (1-78)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||   |||:::|||||||||||||   :::::::|||
Db      54 CTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTC 19


Search completed: May  8, 2003, 23:27:54
Job time : 63 secs
```

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:53:08 ; Search time 207 Seconds
(without alignments)
881.217 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 710428

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113851_7658/app_query.fasta_1.263
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_208_@runat_07052003_113851_7658 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	56	12.7	42	24	AAD27031	Human Secreted epi
c	2	49	11.1	50	8	AAN70336	3'-5' sequence of
	3	47	10.6	51	22	AAI78209	Human silent SNP c
c	4	47	10.6	60	24	ABN39530	Human spliced tran
c	5	46	10.4	60	24	ABN45376	Human spliced tran
c	6	46	10.4	65	24	ABN52611	Mouse spliced tran
	7	46	10.4	66	24	ABN73405	Bovine embryonic g
	8	46	10.4	68	20	AAX35998	Polynucleotide ass
	9	46	10.4	71	21	AAA49245	HCV core protein p
	10	46	10.4	71	21	AAA46147	Plasmid pgHCV/delt
c	11	46	10.4	72	14	AAQ51446	Human FACC intron
	12	46	10.4	79	14	AAQ36160	Mutagenic primer C
c	13	45	10.2	48	22	AAI76614	Human silent SNP c
	14	44.5	10.1	65	24	ABN31870	Rat spliced transc
c	15	44	10.0	50	15	AAQ54486	Excitatory amino a
c	16	44	10.0	50	15	AAQ54487	Excitatory amino a
	17	44	10.0	60	21	AAA50222	Hybrid leptin/tiss
	18	44	10.0	63	21	AAA50221	Hybrid leptin/tiss
	19	44	10.0	75	22	AAD09429	V1Jns-tPA vaccine
c	20	43.5	9.8	65	24	ABN31369	Rat spliced transc
	21	43.5	9.8	76	21	AAA99033	Human TGC839 PCR p
	22	43.5	9.8	80	24	ABL59337	PCR primer for hum
	23	43	9.7	50	22	AAL30844	Human SNP oligonuc
	24	43	9.7	51	22	AAI75717	Human silent SNP c
	25	43	9.7	60	24	ABN38069	Human spliced tran
c	26	43	9.7	60	24	ABN40440	Human spliced tran
c	27	43	9.7	65	24	ABN31099	Rat spliced transc
	28	43	9.7	66	21	AAA49049	DNA encoding tissu
	29	43	9.7	66	21	AAA50218	Modified human tis
	30	43	9.7	66	22	AAC66513	Tissue plasminogen
	31	43	9.7	69	21	AAA50217	Native human tissu
	32	43	9.7	69	21	AAA50219	Modified human tis
	33	43	9.7	69	21	AAA50220	Modified human tis
	34	43	9.7	69	21	AAA09463	tPA leader sequenc
	35	43	9.7	69	24	ABL39945	Signal peptide Tpa
c	36	43	9.7	70	23	ABL58043	Cyclooxygenase-2,
	37	43	9.7	75	24	ABL39944	Signal peptide Tpa
	38	43	9.7	76	21	AAA50338	Synthetic oligonuc
	39	43	9.7	78	17	AAT15476	Human tissue-speci
	40	43	9.7	78	18	AAT85512	Human tissue-speci
	41	43	9.7	78	19	AAV21738	Sense oligomer I u
c	42	43	9.7	78	19	AAV21739	Antisense oligomer
	43	43	9.7	78	22	AAS09203	Synthetic sense ol
c	44	43	9.7	78	22	AAS09204	Synthetic antisens
c	45	43	9.7	78	22	AAD09430	V1Jns-tPA vaccine

ALIGNMENTS

RESULT 1
AAD27031
ID AAD27031 standard; DNA; 42 BP.
XX
AC AAD27031;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 DNA amplifying PCR primer #5.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human; PCR primer; ss.
XX

OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Example 4; Page 89; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC DNA amplifying PCR primer.
XX
SQ Sequence 42 BP; 15 A; 10 C; 9 G; 8 T; 0 other;

Alignment Scores:

Pred. No.:	224	Length:	42
Score:	56.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.67%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x AAD27031 (1-42)

```
Qy      25 LysArgArgProAlaLysAlaTrpSerGly 34
           ||||||||||||||||||||||||||||||
Db      13 AAACGTCGTCCAGCTAAAGCCTGGTCAGGC 42
```

RESULT 2
AAN70336/c
ID AAN70336 standard; DNA; 50 BP.
XX
AC AAN70336;
XX
DT 03-OCT-2002 (updated)
DT 02-APR-1991 (first entry)
XX
DE 3'-5' sequence of probe A11 used to detect hepatitis B virus (HBV)
DE analyte.
XX
KW Hepatitis B virus assay; diagnosis; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /*tag= a
FT /note="Fluorescein label conjugate binding site"
FT misc_feature 21..50
FT /*tag= b
FT /note="probe segment to HBV"
XX
PN EP225807-A.
XX
PD 16-JUN-1987.
XX
PF 10-DEC-1986; 86EP-0309622.
XX
PR 11-DEC-1985; 85US-0807624.
PR 23-DEC-1986; 86US-0945876.
XX
PA (CHIR-) CHIRON CORP.
XX
PI Ureda MS, Warner B, Horn J;
XX
DR WPI; 1987-165004/24.
XX
PT Detecting specific nucleic acid sequences - by incubating sample
PT with labelling reagent set and capturing reagent set.
XX
PS Example; page 17; 31pp; English.
XX
CC In the example a BglII fragment of hepatitis B virus (AAN70325) was
CC detected. 12 sequences, complimentary to different sequences
CC present in HBV were constructed (see AAN70326-n70337). Six of the HBV
CC complementary sequences are joined to a common sequence (A) for
CC complexing with the label conjugate (A'). The other six HBV
CC complementary sequences are joined to a common sequence (B) for
CC complexing with a biotinylated sequence (B') or a third DNA sequence
CC (B'C) for binding to a support (see AAN70338, AAN70339, AAN70340).
CC (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ Sequence 50 BP; 13 A; 8 C; 14 G; 15 T; 0 other;

Alignment Scores:

Pred. No.:	1.56e+03	Length:	50
Score:	49.00	Matches:	8
Percent Similarity:	78.57%	Conservative:	3
Best Local Similarity:	57.14%	Mismatches:	3
Query Match:	11.09%	Indels:	0
DB:	8	Gaps:	0

US-09-599-087B-5 (1-81) x AAN70336 (1-50)

```
Qy      40 CysCysHisArgValProSerProAsnSerThrAsnLeuLys 53
           ||||||||||:::   ||||||   :::   |||||||||:::
Db      45 TGTTGTCATCAAAGGCCTTCACAACACTTGACCAACTTGCAA 4
```

RESULT 3
AAI78209
ID AAI78209 standard; DNA; 51 BP.
XX
AC AAI78209;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:5150.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.

XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 2086; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 14 A; 13 C; 13 G; 11 T; 0 other;

Alignment Scores:

Pred. No.:	2.63e+03	Length:	51
Score:	47.00	Matches:	8
Percent Similarity:	72.73%	Conservative:	0
Best Local Similarity:	72.73%	Mismatches:	3
Query Match:	10.63%	Indels:	0
DB:	22	Gaps:	0

US-09-599-087B-5 (1-81) x AAI78209 (1-51)

```
Qy     39 LeuCysCysHisArgValProSerProAsnSer 49
          ||||||||   |||   ||||||   ||||||
Db     19 TTGTGCTGTGAACGAGCACCGTCAGAGAATTCC 51
```

RESULT 4
ABN39530/c
ID ABN39530 standard; DNA; 60 BP.
XX
AC ABN39530;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:12278.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 12278; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 13 A; 23 C; 9 G; 15 T; 0 other;

Alignment Scores:

Pred. No.:	3.23e+03	Length:	60
Score:	47.00	Matches:	9
Percent Similarity:	71.43%	Conservative:	1
Best Local Similarity:	64.29%	Mismatches:	4
Query Match:	10.63%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN39530 (1-60)

```
Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
          |||:::|||||||   |||||||||         |||||||||
Db     56 TCGTCTGAAGGCCTAAGGAGACCTTGGCTTGGATGGAGTGGG 15
```

RESULT 5
ABN45376/c
ID ABN45376 standard; DNA; 60 BP.
XX
AC ABN45376;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:18124.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX

PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 18124; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 16 C; 21 G; 9 T; 0 other;

Alignment Scores:

Pred. No.:	4.13e+03	Length:	60
Score:	46.00	Matches:	7
Percent Similarity:	63.64%	Conservative:	0
Best Local Similarity:	63.64%	Mismatches:	4
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN45376 (1-60)

```
Qy     62 ProCysLysLeuGluProGluProArgLeuTrp 72
          ||||||   |||   |||   ||||||   |||
Db     55 CCGTGCAGTCTCGTTCCGAGCCCGAGATCCTGG 23
```

RESULT 6
ABN52611/c
ID ABN52611 standard; DNA; 65 BP.
XX
AC ABN52611;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:25359.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 25359; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 20 A; 16 C; 13 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	4.58e+03	Length:	65
Score:	46.00	Matches:	8
Percent Similarity:	78.57%	Conservative:	3
Best Local Similarity:	57.14%	Mismatches:	3
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN52611 (1-65)

```
Qy     38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsn 51
          |||:::|||   ||||||:::||||||   :::|||   |||
Db     61 AGAGTTTGCACTCATAGGATCCCAAGTGGTTCTTCATAAAAT 20
```

RESULT 7
ABN73405
ID ABN73405 standard; cDNA; 66 BP.
XX
AC ABN73405;
XX
DT 03-JUL-2002 (first entry)

XX
DE Bovine embryonic germ (EG) cell cDNA EST 990729a CONTIG 2 #4.
XX
KW Bovine; Bos taurus; EST; expressed sequence tag; totipotence;
KW development; gene; ss.
XX
OS Bos taurus.
XX
PN WO200194550-A2.
XX
PD 13-DEC-2001.
XX
PF 07-JUN-2001; 2001WO-US18576.
XX
PR 07-JUN-2000; 2000US-209874P.
PR 06-JUN-2001; 2001US-0876143.
XX
PA (INFI-) INFIGEN INC.
XX
PI Eilertsen KJ, Pfister-Genskow M, Childs L;
XX
DR WPI; 2002-351289/38.
XX
PT An expressed sequence tag (EST), the expression of which, or its
PT complementary sequence, in a cell identifies the cell as a
PT developmentally competent or incompetent cell -
XX
PS Example 16; Page 165; 584pp; English.
XX
CC The present invention describes an expressed sequence tag (EST), where
CC the EST is an isolated, enriched, or purified nucleic acid sequence
CC representing all or part of a gene, the expression of which, or its
CC complementary sequence, in a cell identifies the cell as a
CC developmentally competent or incompetent cell. Molecules which induce
CC developmental competence in a cell line are useful for inducing
CC totipotence in one or more cells. Molecules which induce developmental
CC incompetence in a cell line are useful for preventing a full term
CC pregnancy in an animal and inhibiting totipotence. The molecules are
CC also useful for treating a disease in an animal by inducing development
CC of one or more cells of the animal into a specific cell type. The
CC present sequence represents a bovine EST which is given in the
CC exemplification of the present invention.
XX
SQ Sequence 66 BP; 4 A; 12 C; 6 G; 44 T; 0 other;

Alignment Scores:

Pred. No.:	4.67e+03	Length:	66
Score:	46.00	Matches:	9
Percent Similarity:	66.67%	Conservative:	3
Best Local Similarity:	50.00%	Mismatches:	6
Query Match:	10.41%	Indels:	0
DB:	24	Gaps:	0

US-09-599-087B-5 (1-81) x ABN73405 (1-66)

```
Qy       3 LeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||      :::||||||   ||||||:::|||   |||   |||   :::|||
Db      10 CTTTTTTATATCTCTTCCTTCCTTTGTGTTTTGTGTCTCTTTTTTTTTCTCTTT 63
```

RESULT 8
AAX35998
ID AAX35998 standard; DNA; 68 BP.
XX
AC AAX35998;
XX
DT 15-JUL-1999 (first entry)
XX
DE Polynucleotide associated with the activation of neutrophils.
XX
KW Neutrophil; differentially expressed; bacteria; activation;
KW nucleic acid probe; amplification primer; hybridisation assay;
KW pathogen; ss.
XX
OS Homo sapiens.
XX
PN WO9920642-A2.
XX
PD 29-APR-1999.
XX
PF 21-OCT-1998; 98WO-US22171.
XX
PR 21-OCT-1997; 97US-0062564.
XX
PA (UYYA) UNIV YALE.
XX
PI Goguen J, Newburger P, Prashar Y, Weissman S, Yamaga S;
PI Yerramilli S;
XX
DR WPI; 1999-288271/24.
XX
PT Unique nucleic acid sequence
XX
PS Claim 1; Page 51; 74pp; English.
XX
CC AAX35979-X36064 represents neutrophil nucleic acid sequences which
CC correspond to mRNA, and therefore to genes. These sequences are
CC differentially expressed upon exposure of the neutrophils to virulent
CC or avirulent bacteria. The nucleic acid sequences can be used as
CC nucleic acid probes to screen cDNA and genomic libraries to select
CC similar sequences; to generate amplification primers; and to identify
CC adjacent and flanking sequences in the genome. The sequences can be
CC used, e.g. in hybridisation assays, to detect activated neutrophils or
CC neutrophils previously exposed to virulent or avirulent bacteria. The
CC sequences can also be used as probes to monitor the expression of at
CC lest one differentially expressed neutrophil gene in a method to
CC identify a therapeutic or prophylactic agent that modulates the response
CC of a neutrophil to a pathogen.
XX
SQ Sequence 68 BP; 17 A; 7 C; 25 G; 17 T; 2 other;

Alignment Scores:

Pred. No.:	4.85e+03	Length:	68
Score:	46.00	Matches:	7
Percent Similarity:	63.64%	Conservative:	0
Best Local Similarity:	63.64%	Mismatches:	4
Query Match:	10.41%	Indels:	0
DB:	20	Gaps:	0

US-09-599-087B-5 (1-81) x AAX35998 (1-68)

```
Qy      31 AlaTrpSerGlyArgArgThrArgLeuCysCys 41
             ||||||   |||||||||         ||||||
Db       2 GCCTGGGANGGGAGGAGAGGAGAGNCGTGCTGT 34
```

RESULT 9
AAA49245
ID AAA49245 standard; DNA; 71 BP.
XX
AC AAA49245;
XX
DT 19-DEC-2000 (first entry)
XX
DE HCV core protein plasmid pgHCVdelta145-154 mutagenic oligonucleotide #12.
XX
KW Hepatitis C; targeted protein production; core protein;
KW mutagenic primer; ss.
XX
OS Synthetic.
XX
PN WO200031127-A2.
XX
PD 02-JUN-2000.
XX
PF 23-NOV-1999; 99WO-GB03898.
XX
PR 26-NOV-1998; 98GB-0025953.

XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Hope G, Mclauchlan J;
XX
DR WPI; 2000-400040/34.
XX
PT Protein comprising a lipid globule targeting sequence consisting of a
PT hepatitis C virus core protein, useful for targeting a protein of
PT interest to lipid globules which are subsequently secreted into animal
PT milk -
XX
PS Example; Page 32; 60pp; English.
XX
CC The present sequence is a mutagenic primer used in the construction of a
CC plasmid containing a mutated version of the hepatitis C virus core
CC protein gene. The core protein has been shown to be localised to lipid
CC droplets within cells. This means that fragments of the protein, in
CC particular those comprising amino acids 125-144 or 161-166, can be used
CC to target proteins to lipid droplets in milk. This provides a simple and
CC efficient means of obtaining proteins of interest, using transgenic
CC animals. Proteins of interest include those implicated in disease and
CC antigenic polypeptides for use as vaccines.
XX
SQ Sequence 71 BP; 13 A; 20 C; 22 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	5.13e+03	Length:	71
Score:	46.00	Matches:	9
Percent Similarity:	68.75%	Conservative:	2
Best Local Similarity:	56.25%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	21	Gaps:	0

US-09-599-087B-5 (1-81) x AAA49245 (1-71)

```
Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66
```

RESULT 10
AAA46147
ID AAA46147 standard; cDNA; 71 BP.
XX
AC AAA46147;
XX
DT 27-OCT-2000 (first entry)
XX
DE Plasmid pgHCV/delta145-154 insertion oligonucleotide.
XX
KW Hepatitis C virus; core protein; HCV; lipid globule; chronic hepatitis;
KW liver disease; ADRP displacement; insertion oligonucleotide; ss.
XX
OS Synthetic.
XX
PN WO200031532-A1.
XX
PD 02-JUN-2000.
XX
PF 23-NOV-1999; 99WO-GB03906.
XX
PR 26-NOV-1998; 98GB-0025951.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Hope G, Mclauchlan J;
XX
DR WPI; 2000-400181/34.
XX
PT Identifying a substance for treating or preventing a viral infection
PT such as hepatitis C virus, comprises determining if the substance
PT disrupts a lipid globule target sequence to lipid globule interaction
PT -
XX
PS Examples; Page 37; 74pp; English.
XX
CC The present sequence is an oligonucleotide which was inserted into
CC a plasmid which contained part of hepatitis c virus core protein gene.
CC The virus causes chronic hepatitis and liver disease in humans. It is
CC thought that it does this by associating with intracellular lipid
CC droplets and downregulating the expression of adipocyte-specific
CC differentiation-related protein (ADRP), which is probably required for
CC the maintenance of lipid droplets. The core protein can be used to
CC identify substances capable of disrupting this interaction, which are
CC thus capable of treating or preventing infection by the hepatitis c virus
CC and its accompanying effects. This oligonucleotide was used in the
CC construction of mutants of the core protein, which were used to determine
CC their efficiency of infection compared to the wild-type.
XX
SQ Sequence 71 BP; 13 A; 20 C; 22 G; 16 T; 0 other;

Alignment Scores:

Pred. No.:	5.13e+03	Length:	71
Score:	46.00	Matches:	9
Percent Similarity:	68.75%	Conservative:	2
Best Local Similarity:	56.25%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	21	Gaps:	0

US-09-599-087B-5 (1-81) x AAA46147 (1-71)

```
Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66
```

RESULT 11
AAQ51446/c
ID AAQ51446 standard; DNA; 72 BP.
XX
AC AAQ51446;
XX
DT 20-MAY-1994 (first entry)
XX
DE Human FACC intron 9 3' region.
XX
KW Fanconi Anemia Group C; FACC; complementing cDNA; varient; diagnosis;
KW open reading frame; Fanconi anemia; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9322435-A.
XX
PD 11-NOV-1993.
XX
PF 27-APR-1993; 93WO-CA00178.
XX
PR 29-APR-1992; 92US-0876285.
PR 21-JUL-1992; 92US-0918313.
PR 15-JAN-1993; 93US-0003963.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
XX
PI Buchwald M, Mathew CG, Strathdee CA, Wevrick R;
XX
DR WPI; 1993-368794/46.
XX
PT Human cDNA which complements Fanconi Anaemia gp. C - used to
PT develop prods. for use in diagnosis, study and therapy of Fanconi
PT Anaemia
XX
PS Claim 1; Page 111; 137pp; English.
XX
CC The sequences given in AAQ51429-54 represent the 5' and 3' splice
CC regions of the introns from the Fanconi Anemia Group C Complementing
CC (FACC) DNA. Three cDNA molecules which are cellular varients of a

```
CC   single cDNA, are transcribed from the FACC gene.  The three cDNAs
CC   each contain an identical open reading frame encoding the FACC
CC   protein.  FACC protein may be used for the diagnosis and study of
CC   Fanconi anemia.  The FACC gene and cDNAs may be used in gene therapy.
XX
SQ   Sequence 72 BP; 12 A; 16 C; 13 G; 31 T; 0 other;

Alignment Scores:
Pred. No.:              5.22e+03        Length:         72
Score:                  46.00           Matches:        7
Percent Similarity:     76.92%          Conservative:   3
Best Local Similarity:  53.85%          Mismatches:     3
Query Match:            10.41%          Indels:         0
DB:                     14              Gaps:           0

US-09-599-087B-5 (1-81) x AAQ51446 (1-72)

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
          ||||||||||:::   :::|||   |||   |||:::|||
Db     71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 12
AAQ36160
ID   AAQ36160 standard; DNA; 79 BP.
XX
AC   AAQ36160;
XX
DT   27-MAY-1993  (first entry)
XX
DE   Mutagenic primer COD1172.
XX
KW   Mutagenesis; site specific; nucleic acid constructs;
KW   restriction site; introduction; removal; ss.
XX
OS   Synthetic.
XX
PN   WO9301282-A.
XX
PD   21-JAN-1993.
XX
PF   01-JUL-1992;   92WO-US05573.
XX
PR   01-JUL-1991;   91US-0724237.
XX
PA   (BERL-) BERLEX LAB INC.
XX
PI   Andrews WH,  Morser MJ,  Vilander LR;
XX
DR   WPI; 1993-045488/05.
XX
PT   Site-specific mutagenesis of nucleic acid constructs - using an
PT   oligo:nucleotide which changes a nucleotide and introduces or
PT   removes a restriction site
XX
PS   Example; Page 48; 87pp; English.
XX
CC   The sequence is that of mutagenic primer COD1172 which was used as part
CC   of a method for site-specific mutagenesis of nucleic acid constructs.
CC   The primer changes a nucleotide and introduces or removes a restriction
CC   site. Using the method multiple mutations can be performed on a single
CC   target sequence simultaneously and more than one such target sequence
CC   can be included on a single construct. The method allows screening of
CC   mutants simply by restriction analysis. The primer was used in the
CC   construction of mutagenesis plasmid pTM3.0-19, it was annealed with
CC   COD1169, COD1170, COD1171 and COD1173 and inserted into the
CC   EcoRI-HindIII sites of pUC18.
XX
SQ   Sequence 79 BP; 10 A; 27 C; 32 G; 10 T; 0 other;

Alignment Scores:
Pred. No.:              5.88e+03        Length:         79
Score:                  46.00           Matches:        8
Percent Similarity:     52.94%          Conservative:   1
Best Local Similarity:  47.06%          Mismatches:     8
Query Match:            10.41%          Indels:         0
DB:                     14              Gaps:           0

US-09-599-087B-5 (1-81) x AAQ36160 (1-79)

Qy     29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValPro 45
          |||      |||:::|||   |||         |||   ||||||   |||
Db     15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCT 65

RESULT 13
AAI76614/c
ID   AAI76614 standard; DNA; 48 BP.
XX
AC   AAI76614;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:3555.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 1139; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 48 BP; 8 A; 15 C; 20 G; 5 T; 0 other;

Alignment Scores:
Pred. No.:              3.97e+03        Length:         48
Score:                  45.00           Matches:        8
Percent Similarity:     76.92%          Conservative:   2
Best Local Similarity:  61.54%          Mismatches:     3
```

```
Query Match:            10.18%      Indels:        0
DB:                     22          Gaps:          0

US-09-599-087B-5 (1-81) x AAI76614 (1-48)

Qy      67 ProGluProArgLeuTrpValValProGlyAlaLeuPro 79
           |||   |||||||||      :::||||||||||:::|||
Db      44 CCCAGTCCCCGACTCCTTGGCCTCCCAGGGGCCATCCCG 6

RESULT 14
ABN31870
ID   ABN31870 standard; DNA; 65 BP.
XX
AC   ABN31870;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:4618.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes
XX
PS   Example 1; SEQ ID 4618; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 20 C; 13 G; 13 T; 0 other;

Alignment Scores:
Pred. No.:              6.63e+03    Length:        65
Score:                  44.50       Matches:       9
Percent Similarity:     70.59%      Conservative:  3
Best Local Similarity:  52.94%      Mismatches:    2
Query Match:            10.07%      Indels:        3
DB:                     24          Gaps:          1

US-09-599-087B-5 (1-81) x ABN31870 (1-65)

Qy      43 ArgValProSerProAsnSer---------ThrAsnLeuLysGlyHisHis 56
           |||:::||||||||||   |||         |||||||||:::   |||:::
Db       1 AGAATTCCAAGCCCTCGTTCCAAAGGTGCAACCAACCTTCAGAGTCACTAT 51

RESULT 15
AAQ54486/c
ID   AAQ54486 standard; cDNA; 50 BP.
XX
AC   AAQ54486;
XX
DT   07-JUL-1994  (first entry)
XX
DE   Excitatory amino acid receptor (EAA4a) coding sequence fragment.
XX
KW   Excitatory amino acid receptor; EAA4a; screening; detection;
KW   identification; CNS; nervous disorders; antibody; oocyte; membrane
KW   protein; ion-channel; therapeutics; ss.
XX
OS   Homo sapiens.
XX
PN   EP578409-A.
XX
PD   12-JAN-1994.
XX
PF   24-JUN-1993;   93EP-0304961.
XX
PR   24-JUN-1992;   92US-0903456.
XX
PA   (ELLI/) ELLIOTT C E.
PA   (KAMB/) KAMBOJ R.
PA   (NUTT/) NUTT S L.
XX
PI   Kamboj R,  Nutt SL,  Elliott CE;
XX
DR   WPI; 1994-010248/02.
XX
PT   Excitatory aminoacid for receptors with kainate binding activity
PT   - used to test ligands for CNS receptor interaction to identify
PT   cpds. useful against CNS disorders
XX
PS   Disclosure; Figure 3b; 37pp; English.
XX
CC   The excitatory amino acid receptor (See AAR45367) can be used to
CC   screen ligands.  The ligands can be assayed for interaction by
CC   incubating them with cells, preferably oocytes, expressing the
CC   receptor or with membrane preparations from these cells, and
CC   assessing any interaction by determining receptor-ligand binding or
CC   ligand-mediated ion channel activation.  This method is useful in
CC   the identification of therapeutics useful to treat CNS disorders in
CC   humans.  Fragments of the receptor are useful for structural
CC   investigations and to raise antibodies.  This fragment of the coding
CC   sequence (AAQ54482) corresponds to bases 1971 to 2020.
XX
SQ   Sequence 50 BP; 11 A; 11 C; 15 G; 13 T; 0 other;

Alignment Scores:
Pred. No.:              5.35e+03    Length:        50
Score:                  44.00       Matches:       7
Percent Similarity:     60.00%      Conservative:  2
Best Local Similarity:  46.67%      Mismatches:    6
Query Match:            9.95%       Indels:        0
DB:                     15          Gaps:          0
```

US-09-599-087B-5 (1-81) x AAQ54486 (1-50)

```
Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||      :::      ||||||   |||   |||:::||||||
Db      46 GGCATGAGCTCAGAACCTTGCTGCATGAGAGCTCCAACTCCAAAC 2
```

Search completed: May 8, 2003, 22:43:14
Job time : 207 secs

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 22:36:33 ; Search time 1229 Seconds
(without alignments)
1918.086 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 463318

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113852_7668/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_1319_@runat_07052003_113852_7668 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		59.5	13.5	76	6	A07238	A07238 Probe for v
2		56	12.7	42	6	AX342232	AX342232 Sequence
3		53	12.0	79	6	I86930	I86930 Sequence 19
4	c	51.5	11.7	69	9	HSU90998	U90998 Homo sapien
5		49	11.1	66	10	MUSTCAYJ	M38677 Mouse T-cel
6		49	11.1	70	10	RNTCRAJ39	Y09178 R.norvegicu
7		48	10.9	76	9	S75385S18	S75420 collagen ty
8		47	10.6	51	6	AX161822	AX161822 Sequence
9		46.5	10.5	48	6	AR077574	AR077574 Sequence
10		46	10.4	71	6	AX025072	AX025072 Sequence
11		46	10.4	71	6	AX025094	AX025094 Sequence
12	c	46	10.4	72	6	I71481	I71481 Sequence 22
13		46	10.4	78	9	HSIGKAB01	X54166 Human BL Ly
14		45.5	10.3	69	6	AR005307	AR005307 Sequence
15		45	10.2	39	6	AR077518	AR077518 Sequence
16	c	45	10.2	48	6	AX160227	AX160227 Sequence
17	c	45	10.2	76	9	AF118405	AF118405 Homo sapi
18		44	10.0	42	6	AR014091	AR014091 Sequence
19	c	44	10.0	50	6	I28909	I28909 Sequence 7
20	c	44	10.0	50	6	I28910	I28910 Sequence 8
21		44	10.0	51	6	AR077572	AR077572 Sequence
22		44	10.0	51	6	AR077575	AR077575 Sequence
23	c	44	10.0	66	6	A28858	A28858 Oligonucleo
24	c	44	10.0	79	5	DUKAETC04	M55135 Peking duck
25	c	43	9.7	46	6	A06605	A06605 Oligonucleo
26	c	43	9.7	46	6	A10221	A10221 adapter. 1/
27		43	9.7	51	6	AX159330	AX159330 Sequence
28		43	9.7	66	6	AX040385	AX040385 Sequence
29		43	9.7	72	9	HSADD2S31	AF000111 Homo sapi
30	c	43	9.7	76	1	MCTRFM	X16759 Mycoplasma
31	c	43	9.7	76	1	MYCTRMF	K00312 Mycoplasma
32		43	9.7	76	6	A07239	A07239 Probe for t
33		43	9.7	78	6	AR000573	AR000573 Sequence
34	c	43	9.7	78	6	AR000574	AR000574 Sequence
35		43	9.7	78	6	AR208886	AR208886 Sequence
36	c	43	9.7	78	6	AR208887	AR208887 Sequence
37		43	9.7	78	6	BD003625	BD003625 A polynuc
38	c	43	9.7	78	6	BD003626	BD003626 A polynuc
39	c	43	9.7	79	6	AR050027	AR050027 Sequence
40	c	42.5	9.6	69	9	HSU91209	U91209 Homo sapien
41	c	42	9.5	30	6	AR022280	AR022280 Sequence
42	c	42	9.5	47	6	AX003724	AX003724 Sequence
43		42	9.5	51	6	AX159329	AX159329 Sequence
44	c	42	9.5	51	6	AX165793	AX165793 Sequence
45		42	9.5	52	9	HS19	X68693 H.sapiens m

ALIGNMENTS

RESULT 1

```
A07238
LOCUS       A07238                    76 bp    DNA     linear   PAT 24-AUG-1993
DEFINITION  Probe for vPA-beta.
ACCESSION   A07238
VERSION     A07238.1  GI:413006
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Baldus,B., Donner,P., Schleuning,W.D., Alagon,A., Boidol,W.,
            Kraetzschmar,J.R., Haendler,B.J. and Langer,G.
  TITLE     Novel thrombolytic
  JOURNAL   Patent: EP 0383417-A 1 22-AUG-1990;
            SCHERING AKTIENGESELLSCHAFT
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       14 a     17 c     24 g     21 t
ORIGIN

Alignment Scores:
Pred. No.:                566           Length:        76
Score:                    59.50         Matches:       14
Percent Similarity:       66.67%        Conservative:  4
Best Local Similarity:    51.85%        Mismatches:    7
Query Match:              13.46%        Indels:        2
DB:                       6             Gaps:          1

US-09-599-087B-5 (1-81) x A07238 (1-76)

Qy      9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePheSer-ThrGluGlyLysArgArgPr 28
          ||||||||||:::|||||||||||||   ::::::||||||       |||   |||:::||
Db      1 CTGTTGTGTGTACTGCTGCTTTGTGGAGCAGTCTTCTCGTTGCCCAGGC---AGGAAACC 57

Qy     28 oAlaLysAlaTrpSerGly 34
          |       |||   |||
Db     58 TACAGGCAATTGGCAAGGG 76

RESULT 2
AX342232
LOCUS       AX342232                  42 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 18 from Patent WO0198497.
ACCESSION   AX342232
VERSION     AX342232.1  GI:18151770
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 18 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR primer corresponding to human SECS-1"
BASE COUNT       15 a     10 c      9 g      8 t
ORIGIN

Alignment Scores:
Pred. No.:                642           Length:        42
Score:                    56.00         Matches:       10
Percent Similarity:       100.00%       Conservative:  0
Best Local Similarity:    100.00%       Mismatches:    0
Query Match:              12.67%        Indels:        0
DB:                       6             Gaps:          0

US-09-599-087B-5 (1-81) x AX342232 (1-42)

Qy     25 LysArgArgProAlaLysAlaTrpSerGly 34
          ||||||||||||||||||||||||||||||
Db     13 AAACGTCGTCCAGCTAAAGCCTGGTCAGGC 42

RESULT 3
I86930
LOCUS       I86930                    79 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 19 from patent US 5702931.
ACCESSION   I86930
VERSION     I86930.1  GI:3206648
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Andrews,W.H., Morser,M.J. and Vilander,L.R.
  TITLE     Mutagenesis methods and compositions
  JOURNAL   Patent: US 5702931-A 19 30-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="unknown"
BASE COUNT       10 a     28 c     31 g     10 t
ORIGIN

Alignment Scores:
Pred. No.:                2.45e+03      Length:        79
Score:                    53.00         Matches:       9
Percent Similarity:       55.00%        Conservative:  2
Best Local Similarity:    45.00%        Mismatches:    9
Query Match:              11.99%        Indels:        0
DB:                       6             Gaps:          0

US-09-599-087B-5 (1-81) x I86930 (1-79)

Qy     29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
          |||      |||:::|||   |||         |||   ||||||   |||   |||:::
Db     15 GCGTCGGGCTGGGCGGGGGCTCGCCAGAGCCGCTGTCGCCACCGTCCACCTTGCCCGAGT 74

RESULT 4
HSU90998/c
LOCUS       HSU90998                  69 bp    mRNA    linear   PRI 15-JUL-1997
DEFINITION  Homo sapiens clone FE1389 T-cell receptor delta chain (TCRDV1J2)
            mRNA, partial cds.
ACCESSION   U90998
VERSION     U90998.1  GI:2239403
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
            Kagnoff,M.F.
  TITLE     The TCR-delta repertoire in human intestine undergoes
            characteristic changes during fetal to adult development
  JOURNAL   J. Immunol. 158 (12), 5632-5641 (1997)
  MEDLINE   97334214
   PUBMED   9190911
REFERENCE   2  (bases 1 to 69)
  AUTHORS   Holtmeier,W., Witthoeft,T., Hennemann,A., Harland,S.W. and
            Kagnoff,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1997) Department of Medicine, University of
            Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="FE1389"
                     /tissue_type="small intestine"
                     /dev_stage="fetus; 20 weeks gestation"
```

```
     gene            <1. .>69
                     /gene="TCRDV1J2"
     CDS             <1. .>69
                     /gene="TCRDV1J2"
                     /note="rearranged; contains CDR3 domain, 11 amino acids
                     (calculation according to: Rock,E., J.Exp.Med.,
                     179:323-328,1994)"
                     /codon_start=1
                     /product="T-cell receptor delta chain"
                     /protein_id="AAC51561.1"
                     /db_xref="GI:2239404"
                     /translation="LEDSAKYFCALGELGDTVTAQLF"
     V_region        <1. .>69
                     /gene="TCRDV1J2"
BASE COUNT       18 a     14 c     19 g     18 t
ORIGIN

Alignment Scores:
Pred. No.:              2.95e+03        Length:         69
Score:                  51.50           Matches:        9
Percent Similarity:     73.33%          Conservative:   2
Best Local Similarity:  60.00%          Mismatches:     3
Query Match:            11.65%          Indels:         1
DB:                     9               Gaps:           1

US-09-599-087B-5 (1-81) x HSU90998 (1-69)

Qy      39 LeuCysCysHisArgValProSer---ProAsnSerThrAsnLeu 52
           |||||||||||||||:::|||      |||   ||||||   :::
Db      63 TTGTGCTGTCACCGTATCCCCCAGTTCCCCAAGAGCACAAAAGTA 19

RESULT 5
MUSTCAYJ
LOCUS       MUSTCAYJ                  66 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T-cell receptor alpha-chain mRNA J-C region, partial cds.
ACCESSION   M38677
VERSION     M38677.1  GI:201214
KEYWORDS    C-region; J-region; T-cell receptor.
SOURCE      Mouse, cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Yague,J., Blackman,M., Born,W., Marrack,P., Kappler,J. and
            Palmer,E.
  TITLE     The structure of V alpha and J alpha segments in the mouse
  JOURNAL   Nucleic Acids Res. 16 (23), 11355-11364 (1988)
  MEDLINE   89083566
   PUBMED   2849763
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .66
                     /partial
                     /codon_start=1
                     /product="T-cell receptor alpha chain"
                     /protein_id="AAA40195.1"
                     /db_xref="GI:201215"
                     /translation="TTASLGKLQFGTGTQVVVTPDI"
BASE COUNT       18 a     16 c     19 g     13 t
ORIGIN

Alignment Scores:
Pred. No.:              4.86e+03        Length:         66
Score:                  49.00           Matches:        8
Percent Similarity:     80.00%          Conservative:   0
Best Local Similarity:  80.00%          Mismatches:     2
Query Match:            11.09%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x MUSTCAYJ (1-66)

Qy      63 CysLysLeuGluProGluProArgLeuTrp 72
           |||   ||||||   |||||||||||||||
Db      23 TGCAGTTTGGAACAGGAACCCAGGTTGTGG 52

RESULT 6
RNTCRAJ39
LOCUS       RNTCRAJ39                 70 bp    mRNA    linear   ROD 06-JAN-1998
DEFINITION  R.norvegicus TCR alpha joining region, clone-library VA8s2F39.
ACCESSION   Y09178
VERSION     Y09178.1  GI:1669729
KEYWORDS    TCR alpha joining region.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Torres-Nagel,N., Deutschlander,A., Herrmann,T., Arden,B. and
            Hunig,T.
  TITLE     Control of TCR V alpha-mediated positive repertoire selection and
            alloreactivity by differential J alpha usage and CDR3 alpha
            composition
  JOURNAL   Int. Immunol. 9 (10), 1441-1452 (1997)
  MEDLINE   98013626
   PUBMED   9352349
REFERENCE   2  (bases 1 to 70)
  AUTHORS   Torres-Nagel,N.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-1996) N.E. Torres-Nagel, Institut of Virology and
            Immunbiology, Department of Immunology, Versbacherstr. 7, 97078
            Wuerzburg, FRG
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Rattus norvegicus"
                     /strain="LEW.1F"
                     /db_xref="taxon:10116"
                     /haplotype="RT1f"
                     /cell_type="peripheral T cell"
                     /clone_lib="VA8s2F39"
                     /dev_stage="adult"
                     /rearranged
     gene            1. .70
                     /gene="TCRAJ"
     CDS             <1. .>70
                     /gene="TCRAJ"
                     /codon_start=1
                     /protein_id="CAA70377.1"
                     /db_xref="GI:1785816"
                     /translation="AVEIATGFGKLQFGTGTQVVVTP"
     J_segment       <1. .>70
                     /gene="TCRAJ"
                     /note="CDR3 region and J alpha region"
                     /evidence=experimental
BASE COUNT       16 a     13 c     24 g     17 t
ORIGIN

Alignment Scores:
Pred. No.:              5.18e+03        Length:         70
Score:                  49.00           Matches:        8
Percent Similarity:     80.00%          Conservative:   0
Best Local Similarity:  80.00%          Mismatches:     2
Query Match:            11.09%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x RNTCRAJ39 (1-70)

Qy      63 CysLysLeuGluProGluProArgLeuTrp 72
           |||   ||||||   |||||||||||||||
Db      32 TGCAGTTTGGAACAGGAACCCAGGTTGTGG 61

RESULT 7
S75385S18
LOCUS       S75385S18                 76 bp    DNA     linear   PRI 02-JUN-2000
```

```
DEFINITION  collagen type VI alpha 1(VI) chain (exons 3-19 and exon/intron
            junctions) [human, Genomic, 76 nt, segment 18 of 18].
ACCESSION   S75420
VERSION     S75420.1  GI:241983
KEYWORDS    .
SEGMENT     18 of 18
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Saitta,B., Wang,Y.M., Renkart,L., Zhang,R.Z., Pan,T.C., Timpl,R.
            and Chu,M.L.
  TITLE     The exon organization of the triple-helical coding regions of the
            human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
  JOURNAL   Genomics 11 (1), 145-153 (1991)
  MEDLINE   92112205
   PUBMED   1765372
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 75420] from the original journal article.
            This sequence comes from Figure 2.
            Map location: chromosome 21.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     mRNA            join(S75385.1:<30. .74,S75388.1:31. .57,S75390.1:31. .57,
                     S75392.1:31. .75,S75394.1:31. .84,S75396.1:31. .93,
                     S75398.1:31. .93,S75400.1:31. .84,S75402.1:31. .66,
                     S75404.1:31. .93,S75406.1:31. .93,S75408.1:31. .93,
                     S75410.1:31. .93,S75412.1:31. .81,S75414.1:31. .66,
                     S75416.1:31. .93,S75418.1:31. .96,31. .>66)
                     /product="collagen type VI alpha 1(VI) chain"
     CDS             join(S75385.1:<30. .74,S75388.1:31. .57,S75390.1:31. .57,
                     S75392.1:31. .75,S75394.1:31. .84,S75396.1:31. .93,
                     S75398.1:31. .93,S75400.1:31. .84,S75402.1:31. .66,
                     S75404.1:31. .93,S75406.1:31. .93,S75408.1:31. .93,
                     S75410.1:31. .93,S75412.1:31. .81,S75414.1:31. .66,
                     S75416.1:31. .93,S75418.1:31. .96,31. .>66)
                     /gene="collagen type VI alpha 1(VI) chain"
                     /note="This sequence comes from Figure 2; conceptual
                     translation differs from the translation provided in the
                     manuscript"
                     /codon_start=1
                     /product="collagen type VI alpha 1(VI) chain"
                     /protein_id="AAB20835.2"
                     /db_xref="GI:8176524"
                     /translation="GRPGDLGPVGYQGMKGEKGSRGEKGSRGPKGYKGEKGKRGIDGV
                     DGVKGEMGYPGLPGCKGSPGFDGIQGPPGPKGDPGAFGLKGEKGEPGADGEAGRPGAR
                     GPSGDEGPAGEPGPPGEKGEAGDEGNPGPDGAPGERGGPGERGPRGTPGTRGPRGDPG
                     EAGPQGDQGREGPVGVPGDPGEAGPIGPKGYRGDEGPPGSEGARGAPGPAGPPGDPGL
                     MGERGEDGPAGNGTEGFPGFPGYPGNRGAPGINGTKGYPGLKGDEGEAGDPGDDNNDI
                     APRGVKGAKGYRGPEGPQGPPGHQGPPGPD"
     exon            31. .>66
                     /number=20
BASE COUNT       15 a     26 c     24 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:              7.04e+03        Length:         76
Score:                  48.00           Matches:        9
Percent Similarity:     59.09%          Conservative:   4
Best Local Similarity:  40.91%          Mismatches:     9
Query Match:            10.86%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x S75385S18 (1-76)

Qy      15 LeuCysPheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
           |||   :::|||:::      |||         |||   |||      ||||||:::|||
Db       6 CTCCTCTACTCCGTTTCTCGGACAGGGACCCCCAGGACACCAAGGACCGCCTGGGCCGGA 65

Qy      35 ArgArg 36
           ||||::
Db      66 CGTAAG 71

RESULT 8
AX161822
LOCUS       AX161822                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5150 from Patent WO0140521.
ACCESSION   AX161822
VERSION     AX161822.1  GI:14543153
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5150 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5149 is other entry)
                     Accession number cg43984006"
BASE COUNT       14 a     13 c     13 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:              5.7e+03         Length:         51
Score:                  47.00           Matches:        8
Percent Similarity:     72.73%          Conservative:   0
Best Local Similarity:  72.73%          Mismatches:     3
Query Match:            10.63%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x AX161822 (1-51)

Qy      39 LeuCysCysHisArgValProSerProAsnSer 49
           |||||||||   |||   ||||||   ||||||
Db      19 TTGTGCTGTGAACGAGCACCGTCAGAGAATTCC 51

RESULT 9
AR077574
LOCUS       AR077574                 48 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 289 from patent US 5962255.
ACCESSION   AR077574
VERSION     AR077574.1  GI:10004320
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Griffiths,A.David., Williams,S.Cameron., Waterhouse,P.Michael.,
            Nissim,A., Winter,G.Paul., Johnson,K.Stuart. and
            Smith,A.John.Hammond.
  TITLE     Methods for producing recombinant vectors
  JOURNAL   Patent: US 5962255-A 289 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT        8 a     18 c     11 g     11 t
ORIGIN

Alignment Scores:
Pred. No.:              5.95e+03        Length:         48
Score:                  46.50           Matches:        9
Percent Similarity:     66.67%          Conservative:   3
Best Local Similarity:  50.00%          Mismatches:     3
Query Match:            10.52%          Indels:         3
```

```
DB:                 6           Gaps:          1

US-09-599-087B-5 (1-81) x AR077574 (1-48)

Qy     32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSer 49
          ||||||  ||||||   :::   ||||||          |||:::|||:::|||
Db      3 TGGTCCCTCCGCCGAATACAGTGCTGCTGT---------CCCACGCCTGACAGT 47

RESULT 10
AX025072
LOCUS       AX025072                71 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 18 from Patent WO0031127.
ACCESSION   AX025072
VERSION     AX025072.1  GI:10184987
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Hope,G. and Mclauchlan,J.
  TITLE     Protein fragments for use in protein targeting
  JOURNAL   Patent: WO 0031127-A 18 02-JUN-2000;
            MEDICAL RES COUNCIL (GB) ; HOPE GRAHAM (GB) ; MCLAUCHLAN JOHN (GB)
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide used in mutagenesis"
BASE COUNT       13 a     20 c     22 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              1.01e+04        Length:         71
Score:                  46.00           Matches:        9
Percent Similarity:     68.75%          Conservative:   2
Best Local Similarity:  56.25%          Mismatches:     5
Query Match:            10.41%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x AX025072 (1-71)

Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66

RESULT 11
AX025094
LOCUS       AX025094                71 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 18 from Patent WO0031532.
ACCESSION   AX025094
VERSION     AX025094.1  GI:10185008
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Hope,G. and Mclauchlan,J.
  TITLE     Methods of identifying viral therapeutics
  JOURNAL   Patent: WO 0031532-A 18 02-JUN-2000;
            MEDICAL RES COUNCIL (GB) ; HOPE GRAHAM (GB) ; MCLAUCHLAN JOHN (GB)
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide used in mutagenesis"
BASE COUNT       13 a     20 c     22 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              1.01e+04        Length:         71
Score:                  46.00           Matches:        9
Percent Similarity:     68.75%          Conservative:   2
Best Local Similarity:  56.25%          Mismatches:     5
Query Match:            10.41%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x AX025094 (1-71)

Qy     25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          :::||||||||||:::         ||||||||||||      ||||||
Db     19 CGTCGGCGCCCCTCTTGTCCGGGTTCTGGAAGACGGTGTGAACTATGC 66

RESULT 12
I71481/c
LOCUS       I71481                  72 bp    DNA     linear   PAT 03-APR-1998
DEFINITION  Sequence 22 from patent US 5681942.
ACCESSION   I71481
VERSION     I71481.1  GI:3007616
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Buchwald,M., Strathdee,C.A., Wevrick,R. and Mathew,C.George.Porter.
  TITLE     Fanconi Anemia Type C gene
  JOURNAL   Patent: US 5681942-A 22 28-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="unknown"
BASE COUNT       12 a     16 c     13 g     31 t
ORIGIN

Alignment Scores:
Pred. No.:              1.03e+04        Length:         72
Score:                  46.00           Matches:        7
Percent Similarity:     76.92%          Conservative:   3
Best Local Similarity:  53.85%          Mismatches:     3
Query Match:            10.41%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x I71481 (1-72)

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLys 53
          ||||||||||:::   :::|||   |||   |||:::|||
Db     71 TGCCACAGGATGGAAAATCCAAAGAGCATGAACATTAAG 33

RESULT 13
HSIGKAB01
LOCUS       HSIGKAB01               78 bp    mRNA    linear   PRI 04-AUG-1992
DEFINITION  Human BL Ly66 mRNA for abnormal Ig kappa light chain (1).
ACCESSION   X54166
VERSION     X54166.1  GI:33218
KEYWORDS    Ig kappa light chain; Ig light chain; immunoglobulin.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   COGNE,M.C.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1990) Cogne M.C.C., CNRS URA 1172, Lab.
            Immunologie Moleculaire, Faculte des Sciences, F-86022 Poitiers,
            france
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Cogne,M., Mounir,S., Mahdi,T., Preud'homme,J.L., Nau,F. and
            Guglielmi,P.
  TITLE     Production of an abnormal mu chain with a shortened VHIV subgroup
            variable region in a Burkitt's lymphoma cell line
  JOURNAL   Mol. Immunol. 27 (9), 929-934 (1990)
  MEDLINE   91015095
   PUBMED   2120579
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="Ly66"
```

```
                     /cell_type="B-cell"
                     /tissue_type="Burkitt lymphoma"
     mRNA            1. .78
                     /note="abnormal Ig L chain 0.85kb transcript"
                     /evidence=experimental
     exon            14. .68
                     /note="V kappa 1 leader peptide exon"
                     /number=1
     exon            69. .>78
                     /note="C kappa exon, corresponds to Km3  allele sequence"
                     /number=2
BASE COUNT       13 a     24 c     23 g     18 t
ORIGIN

Alignment Scores:
Pred. No.:              1.12e+04        Length:         78
Score:                  46.00           Matches:        10
Percent Similarity:     76.47%          Conservative:   3
Best Local Similarity:  58.82%          Mismatches:     4
Query Match:            10.41%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x HSIGKAB01 (1-78)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPhe 17
          |||   :::   ||||||:::   ||||||   :::|||||||||||||||
Db     14 ATGGACATGAGAGTCCTCGCTCAGCTCCTGGGGCTCCTGCTGCTCTGTTTC 64

RESULT 14
AR005307
LOCUS       AR005307                69 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 37 from patent US 5747660.
ACCESSION   AR005307
VERSION     AR005307.1  GI:3966186
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Orlicky,D.J.
  TITLE     Nucleic acid encoding prostaglandin F.sub.2.alpha. receptor
            regulatory protein
  JOURNAL   Patent: US 5747660-A 37 05-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="unknown"
BASE COUNT        6 a     25 c     31 g      7 t
ORIGIN

Alignment Scores:
Pred. No.:              1.1e+04         Length:         69
Score:                  45.50           Matches:        10
Percent Similarity:     54.55%          Conservative:   2
Best Local Similarity:  45.45%          Mismatches:     7
Query Match:            10.29%          Indels:         3
DB:                     6               Gaps:           1

US-09-599-087B-5 (1-81) x AR005307 (1-69)

Qy     25 LysArgArgProAla---------LysAlaTrpSerGlyArgArgThrArgLeuCysCys 41
          :::|||||||   |||         :::|||||||      ||||||         ||||||
Db      3 AGGCGGAGAGTCGCTCCCGCCGGCCGAGCATGGGGCGCCCGGCGCCGAGGCCGCTGCTGC 62

Qy     42 HisArg 43
            |||
Db     63 TGGCGC 68

RESULT 15
AR077518
LOCUS       AR077518                39 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 233 from patent US 5962255.
ACCESSION   AR077518
VERSION     AR077518.1  GI:10004264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Griffiths,A.David., Williams,S.Cameron., Waterhouse,P.Michael.,
            Nissim,A., Winter,G.Paul., Johnson,K.Stuart. and
            Smith,A.John.Hammond.
  TITLE     Methods for producing recombinant vectors
  JOURNAL   Patent: US 5962255-A 233 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT        7 a     13 c     10 g      9 t
ORIGIN

Alignment Scores:
Pred. No.:              6.61e+03        Length:         39
Score:                  45.00           Matches:        8
Percent Similarity:     66.67%          Conservative:   0
Best Local Similarity:  66.67%          Mismatches:     4
Query Match:            10.18%          Indels:         0
DB:                     6               Gaps:           0

US-09-599-087B-5 (1-81) x AR077518 (1-39)

Qy     32 TrpSerGlyArgArgThrArgLeuCysCysHisArg 43
          ||||||   ||||||         ||||||||||||
Db      3 TGGTCCCTCCGCCGAATACCAGGCTGCTGTCATAGG 38


Search completed: May  8, 2003, 23:26:38
Job time : 1229 secs
```

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:46:33 ; Search time 21 Seconds
(without alignments)
354.956 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 44737

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61.5	13.9	76	9	US-10-178-213-128	Sequence 128, App
2	61.5	13.9	76	9	US-10-178-213-131	Sequence 131, App
3	58.5	13.2	77	9	US-10-097-065-142	Sequence 142, App
4	58	13.1	53	9	US-09-981-876-255	Sequence 255, App
5	58	13.1	53	9	US-09-148-545-255	Sequence 255, App
6	55.5	12.6	52	9	US-09-974-879-233	Sequence 233, App
7	55	12.4	80	9	US-10-097-065-432	Sequence 432, App
8	53.5	12.1	44	10	US-09-864-761-46147	Sequence 46147, A
9	53	12.0	61	10	US-09-815-242-4938	Sequence 4938, Ap
10	52.5	11.9	59	9	US-09-809-391-680	Sequence 680, App
11	52.5	11.9	62	10	US-09-739-254-108	Sequence 108, App
12	52.5	11.9	62	10	US-09-904-615-108	Sequence 108, App
13	51.5	11.7	61	9	US-10-097-065-610	Sequence 610, App
14	51	11.5	35	9	US-09-291-925A-1	Sequence 1, Appli
15	51	11.5	47	9	US-09-809-391-523	Sequence 523, App
16	51	11.5	80	9	US-10-178-213-89	Sequence 89, Appl
17	49.5	11.2	60	10	US-09-864-761-40923	Sequence 40923, A
18	49.5	11.2	73	9	US-10-079-854-174	Sequence 174, App
19	49.5	11.2	73	10	US-09-764-878-174	Sequence 174, App
20	48	10.9	61	9	US-10-259-520-9	Sequence 9, Appli
21	48	10.9	61	10	US-09-263-811-9	Sequence 9, Appli
22	48	10.9	62	9	US-10-091-504-839	Sequence 839, App
23	48	10.9	62	10	US-09-764-869-839	Sequence 839, App
24	48	10.9	66	9	US-10-163-547-20	Sequence 20, Appl
25	47.5	10.7	44	10	US-09-864-761-40752	Sequence 40752, A
26	47.5	10.7	69	9	US-09-910-009A-65	Sequence 65, Appl
27	47.5	10.7	69	9	US-09-910-009A-219	Sequence 219, App
28	47.5	10.7	79	9	US-10-091-504-889	Sequence 889, App
29	47.5	10.7	79	10	US-09-764-869-889	Sequence 889, App
30	47	10.6	32	9	US-09-291-925A-6	Sequence 6, Appli
31	47	10.6	72	9	US-09-898-659-29	Sequence 29, Appl
32	46.5	10.5	32	9	US-09-983-802-207	Sequence 207, App
33	46.5	10.5	57	10	US-09-802-094-6	Sequence 6, Appli
34	46.5	10.5	64	9	US-09-736-457-1830	Sequence 1830, Ap
35	46.5	10.5	64	9	US-09-902-941-1830	Sequence 1830, Ap
36	46.5	10.5	64	9	US-09-849-626-1830	Sequence 1830, Ap
37	46.5	10.5	64	9	US-10-017-754-1830	Sequence 1830, Ap
38	46.5	10.5	68	9	US-10-072-349-153	Sequence 153, App
39	46.5	10.5	68	10	US-09-764-855-153	Sequence 153, App
40	46.5	10.5	71	10	US-09-925-300-1861	Sequence 1861, Ap
41	46.5	10.5	73	8	US-08-424-550B-32	Sequence 32, Appl
42	46.5	10.5	76	10	US-09-925-301-1187	Sequence 1187, Ap
43	46	10.4	40	10	US-09-864-761-45258	Sequence 45258, A
44	46	10.4	61	9	US-10-012-542-486	Sequence 486, App
45	46	10.4	61	10	US-09-867-550-1642	Sequence 1642, Ap

ALIGNMENTS

RESULT 1
US-10-178-213-128
; Sequence 128, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Nicotiana benthamiana
US-10-178-213-128

Query Match 13.9%; Score 61.5; DB 9; Length 76;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 14; Conservative 9; Mismatches 10; Indels 9; Gaps 3;

```
Qy      10 LCILLLCF------SIFSTEGKR--RPAKAWSGR-RTRLCCH 42
           :|::| ||       :  || |   :|:|:|:|:   | | |
Db      11 ICVVLFCFLLMFASEVQITEAKHCGKPSKSWNGKCFPRKCNH 52
```

RESULT 2
US-10-178-213-131
; Sequence 131, Application US/10178213

```
; Publication No. US20030041348A1
; GENERAL INFORMATION:
;  APPLICANT: Simmons, Carl R.
;  APPLICANT:  Navarro Acevedo, Pedro A.
;  APPLICANT:  Harvell, Leslie
;  APPLICANT:  Cahoon, Rebecca
;  APPLICANT:  McCutchen, Billy Fred
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Herrmann, Rafael
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Defensin Polynucleotides and Methods of
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: 35718/246703
;  CURRENT APPLICATION NUMBER: US/10/178,213
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: 60/300,152
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: 60/300,241
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131
;   LENGTH: 76
;   TYPE: PRT
;   ORGANISM: Nicotiana benthamiana
US-10-178-213-131

  Query Match             13.9%;  Score 61.5;  DB 9;  Length 76;
  Best Local Similarity   33.3%;  Pred. No. 14;
  Matches   14;  Conservative    9;  Mismatches   10;  Indels    9;  Gaps    3;

Qy      10 LCILLLCF------SIFSTEGKR--RPAKAWSGR-RTRLCCH 42
           :|::| ||       :  || |   :|:|:|:|:   | | |
Db      11 ICVVLFCFLLMFASEVQITEAKHCGKPSKSWNGKCFPRKCNH 52


RESULT 3
US-10-097-065-142
; Sequence 142, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, Paul A. et al.
;  TITLE OF INVENTION: 110 Human Secreted Proteins
;  FILE REFERENCE: PZ021P1
;  CURRENT APPLICATION NUMBER: US/10/097,065
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: PCT/US98/27059
;  PRIOR FILING DATE: 1998-12-17
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,007
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,057
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,006
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,369
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,367
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,368
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,169
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,008
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,365
;  PRIOR FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 142
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (77)
;   OTHER INFORMATION: Xaa equals stop translation
US-10-097-065-142

  Query Match             13.2%;  Score 58.5;  DB 9;  Length 77;
  Best Local Similarity   32.8%;  Pred. No. 30;
  Matches   19;  Conservative    5;  Mismatches   17;  Indels   17;  Gaps    2;

Qy       4 LVLSSLLCILLLCF--------------SIFSTEGKR--RPAKAWSGRRTRLCCHRV 44
             ||| |:  |||:                 |:  | |:|  || :   |   || | ||
Db      11 LVLGSVFMILLIIVYWDSAAPRTSTCTRPSLGRTRGRRCPRPGRTGQGAHGRLRCRRV 68


RESULT 4
US-09-981-876-255
; Sequence 255, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 70 Human Secreted Proteins
;  FILE REFERENCE: PZ001P1
;  CURRENT APPLICATION NUMBER: US/09/981,876
;  CURRENT FILING DATE:  2001-10-19
;  PRIOR APPLICATION NUMBER: 09/148,545
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 60/040,162
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,333
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/038,621
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,161
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,626
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,334
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,336
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,163
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/047,615
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,600
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,597
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,502
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,633
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,583
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,617
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,618
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,503
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,592
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,581
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,584
```

; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043;671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/057,761
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884

```
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 280
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
;   LENGTH: 53

  Query Match             13.1%;  Score 58;  DB 9;  Length 53;
  Best Local Similarity   50.0%;  Pred. No. 23;
  Matches   12;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 LLVLSSLLCILLLCFSIFSTEGKR 26
           ||:|:||  : | |: :| |:|||
Db      19 LLLLASLAPLWLYCWRMFPTKGKR 42


RESULT 5
US-09-148-545-255
; Sequence 255, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 70 Human Secreted Proteins
;  FILE REFERENCE: PZ001P1
;  CURRENT APPLICATION NUMBER: US/09/148,545
;  CURRENT FILING DATE: 1998-09-04
;  EARLIER APPLICATION NUMBER: PCT/US98/04482
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,161
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,674
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,669
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,312
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,313
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,672
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,315
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/056,886
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,877
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,889
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,893
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,630
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,878
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,662
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,872
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,882
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,637
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,903
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,888
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,879
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,880
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,894
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,911
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,636
```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 53

Query Match 13.1%; Score 58; DB 9; Length 53;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 12; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LLVLSSLLCILLLCFSIFSTEGKR 26
||:|:|| : | |: :| |:|||
Db 19 LLLLASLAPLWLYCWRMFPTKGKR 42

RESULT 6
US-09-974-879-233
; Sequence 233, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: PZ020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 233
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-879-233

Query Match 12.6%; Score 55.5; DB 9; Length 52;
Best Local Similarity 34.0%; Pred. No. 40;
Matches 17; Conservative 4; Mismatches 24; Indels 5; Gaps 2;

Qy 8 SLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNL--KGH 55
| ||:| | : | | : || |: || | | |||
Db 6 STLCVLSFCSVVCSVEASCQSTVQWGGAAARV---GVPFDWSRNEQGKGH 52

RESULT 7
US-10-097-065-432

```
; Sequence 432, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, Paul A. et al.
;  TITLE OF INVENTION: 110 Human Secreted Proteins
;  FILE REFERENCE: PZ021P1
;  CURRENT APPLICATION NUMBER: US/10/097,065
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: PCT/US98/27059
;  PRIOR FILING DATE: 1998-12-17
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,007
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,057
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,006
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,369
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,367
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,368
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,169
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,008
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,365
;  PRIOR FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 432
;   LENGTH: 80
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-097-065-432

  Query Match             12.4%;  Score 55;  DB 9;  Length 80;
  Best Local Similarity   33.3%;  Pred. No. 73;
  Matches   16;  Conservative    7;  Mismatches   23;  Indels    2;  Gaps    1;

Qy       8 SLLCILLL--CFSIFSTEGKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 53
           |||. : ||  | : ||  |   |  | |  |    | ||  |: :: ::
Db      15 SLLLVQLLTPCSAQFSVLGPSGPILAMVGEDADLPCHLFPTMSAETME 62


RESULT 8
US-09-864-761-46147
; Sequence 46147, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46147
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC023890.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
;   OTHER INFORMATION: SWISSPROT HIT: P45816, EVALUE 4.80e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BE545503.1, EVALUE 7.00e-19
US-09-864-761-46147

  Query Match             12.1%;  Score 53.5;  DB 10;  Length 44;
  Best Local Similarity   37.1%;  Pred. No. 55;
  Matches   13;  Conservative    6;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      45 PSPN-STNLKGHHVRLCKPCKLEPEPRLWVVPGAL 78
           |||: | :|: |  ||| |  |  :    ::|  |
Db       4 PSPSPSPSLRCHSARLCMPMTLRTQTNSALMPMTL 38


RESULT 9
US-09-815-242-4938
; Sequence 4938, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
```

```
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4938
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-815-242-4938

  Query Match             12.0%;  Score 53;  DB 10;  Length 61;
  Best Local Similarity   40.0%;  Pred. No. 88;
  Matches   14;  Conservative    4;  Mismatches   15;  Indels    2;  Gaps    1;

Qy      26 RRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLC 60
           :|||| :   ||  | |   |:|   | |  |:|
Db      11 KRPAKHSTQAYTR--CERCGRPHSVYRKFHLCRIC 43


RESULT 10
US-09-809-391-680
; Sequence 680, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P2
;  CURRENT APPLICATION NUMBER: US/09/809,391
;  CURRENT FILING DATE:  2001-03-16
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 761
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 680
;   LENGTH: 59
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-809-391-680

  Query Match             11.9%;  Score 52.5;  DB 9;  Length 59;
  Best Local Similarity   33.3%;  Pred. No. 96;
  Matches   14;  Conservative    7;  Mismatches   18;  Indels    3;  Gaps    2;

Qy      27 RPAKAWSGRRTRLCCHR-VPSPNSTNLKGHHVRLCKPCKLEP 67
           ||:: |   || :  ||   :  |:  |   ::|   || ||
Db      13 RPSRGWRSSRTSVSRHRDTENTRSSRSKTGSLQLI--CKSEP 52


RESULT 11
US-09-739-254-108
; Sequence 108, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/09/739,254
;  CURRENT FILING DATE: 2000-12-19
;  EARLIER APPLICATION NUMBER: 09/511,554
;  EARLIER FILING DATE: 2000-02-23
;  EARLIER APPLICATION NUMBER: PCT/US99/19330
;  EARLIER FILING DATE: 1999-08-24
;  EARLIER APPLICATION NUMBER: 60/097,917
;  EARLIER FILING DATE: 1998-08-25
;  EARLIER APPLICATION NUMBER: 60/098,634
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 108
;   LENGTH: 62
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (62)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-739-254-108

  Query Match             11.9%;  Score 52.5;  DB 10;  Length 62;
  Best Local Similarity   29.3%;  Pred. No. 1e+02;
  Matches   17;  Conservative   11;  Mismatches   23;  Indels    7;  Gaps    3;

Qy      10 LCILLLCFSIFS-TEGKRR----PAKAWSGRRTRLCCHRVPSP--NSTNLKGHHVRLC 60
           :|:|||  : |  | |  |     : ||: |   :   : |||  ::  |:|    :|
Db       1 MCLLLLWLTTFQRTSGALRRGGLSSPAWAMRSPSVYSTQTPSPMMSTGTLRGLSGAMC 58


RESULT 12
US-09-904-615-108
; Sequence 108, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/09/904,615
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/511,554
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/097,917
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 60/098,634
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 108
;   LENGTH: 62
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (62)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-904-615-108

  Query Match             11.9%;  Score 52.5;  DB 10;  Length 62;
  Best Local Similarity   29.3%;  Pred. No. 1e+02;
  Matches   17;  Conservative   11;  Mismatches   23;  Indels    7;  Gaps    3;

Qy      10 LCILLLCFSIFS-TEGKRR----PAKAWSGRRTRLCCHRVPSP--NSTNLKGHHVRLC 60
           :|:|||  : |  | |  |     : ||: |   :   : |||  ::  |:|    :|
Db       1 MCLLLLWLTTFQRTSGALRRGGLSSPAWAMRSPSVYSTQTPSPMMSTGTLRGLSGAMC 58


RESULT 13
US-10-097-065-610
; Sequence 610, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, Paul A. et al.
;  TITLE OF INVENTION: 110 Human Secreted Proteins
;  FILE REFERENCE: PZ021P1
;  CURRENT APPLICATION NUMBER: US/10/097,065
```

```
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: PCT/US98/27059
;  PRIOR FILING DATE: 1998-12-17
;  PRIOR APPLICATION NUMBER: 60/070,923
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,007
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,057
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,006
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,369
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,367
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,368
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,169
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/068,053
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,064
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,054
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,008
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/068,365
;  PRIOR FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 610
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-097-065-610

  Query Match           11.7%;  Score 51.5;  DB 9;  Length 61;
  Best Local Similarity  30.4%;  Pred. No. 1.3e+02;
  Matches   21;  Conservative    8;  Mismatches   15;  Indels   25;  Gaps    4;

Qy      14 LLCFSIFSTEGKRRPAKA-WSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPR-- 70
           ||   ||:|    ||: | :|| ||     | |:|                    |:||
Db      14 LLALQIFNT--TFRPSFAFFSGHRTLFFGVRSPNP------------------PKPRIF 52

Qy      71 -LWVVPGAL 78
            :|::  ||
Db      53 LIWLIAVAL 61


RESULT 14
US-09-291-925A-1
; Sequence 1, Application US/09291925A
; Patent No. US20020155525A1
; GENERAL INFORMATION:
;  APPLICANT:  Ashkenazi, A.
;  APPLICANT:  Berman, P.
;  APPLICANT:  Brousseau, D.
;  APPLICANT:  Etcheverry, T.
;  TITLE OF INVENTION:  SECRETION OF GLYCOSYLATION MUTANTS
;  FILE REFERENCE:  P1055R1
;  CURRENT APPLICATION NUMBER: US/09/291,925A
;  CURRENT FILING DATE:  1999-04-14
;  PRIOR APPLICATION NUMBER:  US 60/082,002
;  PRIOR FILING DATE:  1998-04-16
;  PRIOR APPLICATION NUMBER:  US 60/123,522
;  PRIOR FILING DATE:  1999-03-08
;  NUMBER OF SEQ ID NOS:  13
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-291-925A-1

  Query Match           11.5%;  Score 51;  DB 9;  Length 35;
  Best Local Similarity  34.4%;  Pred. No. 79;
  Matches   11;  Conservative    7;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       5 VLSSLLCILLLCFSIFSTEGKRRPAKAWSGRR 36
           :|  | |:|||| ::| :  :   |:   | |
Db       4 MLRGLCCVLLLCGAVFVSPSQEIHARFRRGAR 35


RESULT 15
US-09-809-391-523
; Sequence 523, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P2
;  CURRENT APPLICATION NUMBER: US/09/809,391
;  CURRENT FILING DATE:  2001-03-16
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 761
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 523
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (17)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (28)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-809-391-523

  Query Match           11.5%;  Score 51;  DB 9;  Length 47;
  Best Local Similarity  46.2%;  Pred. No. 1.1e+02;
  Matches   12;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       3 LLVLSSLLCILLLCFSIFSTEGKRRP 28
           || | | ||:||   |   |  |::|
Db      22 LLALGSXLCLLLCLPSFNKTRRKQKP 47


Search completed: May  8, 2003, 21:49:15
Job time : 23 secs
```

GenCore version 5.1.5

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:45:48 ; Search time 18 Seconds
(without alignments)
132.403 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 46248

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52.5	11.9	59	4	US-09-149-476-680	Sequence 680, App
2	52.5	11.9	74	1	US-08-543-238-5	Sequence 5, Appli
3	52.5	11.9	74	1	US-08-420-526-5	Sequence 5, Appli
4	51	11.5	42	4	US-08-974-549A-174	Sequence 174, App
5	51	11.5	47	4	US-09-149-476-523	Sequence 523, App
6	50	11.3	56	2	US-08-637-759B-123	Sequence 123, App
7	50	11.3	56	3	US-08-871-355A-123	Sequence 123, App
8	50	11.3	56	4	US-09-201-945-123	Sequence 123, App
9	50	11.3	77	2	US-08-726-306A-174	Sequence 174, App
10	49.5	11.2	77	1	US-08-443-568B-10	Sequence 10, Appl
11	49.5	11.2	77	5	PCT-US94-06997-10	Sequence 10, Appl
12	48	10.9	36	6	5256770-8	Patent No. 5256770
13	48	10.9	40	6	5244676-2	Patent No. 5244676
14	48	10.9	61	3	US-08-743-975-9	Sequence 9, Appli
15	47.5	10.7	68	1	US-07-689-693B-11	Sequence 11, Appl
16	47	10.6	35	1	US-07-609-510B-1	Sequence 1, Appli
17	47	10.6	35	1	US-08-427-640-10	Sequence 10, Appl
18	46.5	10.5	32	4	US-09-227-357-207	Sequence 207, App
19	46.5	10.5	45	4	US-08-891-271-8	Sequence 8, Appli
20	46.5	10.5	69	4	US-09-228-302-7	Sequence 7, Appli
21	46.5	10.5	73	4	US-08-469-260A-32	Sequence 32, Appl
22	46	10.4	32	3	US-08-792-832A-48	Sequence 48, Appl
23	46	10.4	66	4	US-09-383-586-15	Sequence 15, Appl
24	45.5	10.3	80	4	US-09-188-930-150	Sequence 150, App
25	45.5	10.3	80	4	US-09-188-930-300	Sequence 300, App
26	45	10.2	28	3	US-08-848-580-10	Sequence 10, Appl
27	45	10.2	28	5	PCT-US92-04537-6	Sequence 6, Appli
28	45	10.2	35	1	US-08-282-030-11	Sequence 11, Appl
29	45	10.2	35	5	PCT-US95-10219-11	Sequence 11, Appl
30	45	10.2	75	2	US-08-783-395-2	Sequence 2, Appli
31	45	10.2	75	2	US-08-924-838-8	Sequence 8, Appli
32	44.5	10.1	38	3	US-08-504-538A-17	Sequence 17, Appl
33	44.5	10.1	38	4	US-08-630-052-17	Sequence 17, Appl
34	44.5	10.1	38	5	PCT-US95-09307-17	Sequence 17, Appl
35	44.5	10.1	46	4	US-08-632-511A-5	Sequence 5, Appli
36	44.5	10.1	46	4	US-09-091-590A-11	Sequence 11, Appl
37	44.5	10.1	46	4	US-09-488-200-5	Sequence 5, Appli
38	44.5	10.1	70	4	US-09-006-428A-15	Sequence 15, Appl
39	44	10.0	52	3	US-08-851-843A-184	Sequence 184, App
40	44	10.0	52	4	US-08-974-549A-303	Sequence 303, App
41	44	10.0	52	4	US-08-854-050-184	Sequence 184, App
42	44	10.0	52	4	US-09-430-323-184	Sequence 184, App
43	44	10.0	67	1	US-08-471-780C-126	Sequence 126, App
44	44	10.0	67	1	US-08-467-282B-126	Sequence 126, App
45	44	10.0	67	2	US-08-471-282A-126	Sequence 126, App

ALIGNMENTS

RESULT 1
US-09-149-476-680
; Sequence 680, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22

```
;  EARLIER APPLICATION NUMBER: 60/057,669
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/049,610
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/061,060
;  EARLIER FILING DATE: 1997-10-02

  Query Match             11.9%;  Score 52.5;  DB 4;  Length 59;
  Best Local Similarity   33.3%;  Pred. No. 11;
  Matches   14;  Conservative    7;  Mismatches   18;  Indels    3;  Gaps    2;

Qy      27 RPAKAWSGRRTRLCCHR-VPSPNSTNLKGHHVRLCKPCKLEP 67
           ||:: |   || :  ||   :  |:  |   ::|   || ||
Db      13 RPSRGWRSSRTSVSRHRDTENTRSSRSKTGSLQLI--CKSEP 52


RESULT 2
US-08-543-238-5
; Sequence 5, Application US/08543238
; Patent No. 5607919
;  GENERAL INFORMATION:
;    APPLICANT:  Bojsen, Kirsten
;    APPLICANT:  Kragh, Karsten M.
;    APPLICANT:  Mikkelsen, Jorn D.
;    APPLICANT:  Nielsen, Klaus K.
;    TITLE OF INVENTION:  Anti-Microbial Proteins
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,238
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marcus-Wyner, Lynn
;      REGISTRATION NUMBER:  34,869
;      REFERENCE/DOCKET NUMBER:  137-1078/MA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3588
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  74 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-543-238-5

  Query Match             11.9%;  Score 52.5;  DB 1;  Length 74;
  Best Local Similarity   25.4%;  Pred. No. 15;
  Matches   17;  Conservative   10;  Mismatches   25;  Indels   15;  Gaps    2;

Qy      12 ILLLCFSIFSTEGK----------RRPAKAWSG-----RRTRLCCHRVPSPNSTNLKGHH 56
           :|||   :|:::|           |:|:  :||        :  |:|   ||   | |
Db       8 LLLLLLFVFASELNMVAEVQGATCRKPSMYFSGACFSDTNCQKACNREDWPNGKCLVGFK 67

Qy      57 VRLCKPC 63
             :||
Db      68 CECQRPC 74


RESULT 3
US-08-420-526-5
; Sequence 5, Application US/08420526
; Patent No. 5608151
;  GENERAL INFORMATION:
;    APPLICANT:  Bojsen, Kirsten
;    APPLICANT:  Kragh, Karsten M.
;    APPLICANT:  Mikkelsen, Jorn D.
;    APPLICANT:  Nielsen, Klaus K.
;    TITLE OF INVENTION:  Anti-Microbial Proteins
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/420,526
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marcus-Wyner, Lynn
;      REGISTRATION NUMBER:  34,869
;      REFERENCE/DOCKET NUMBER:  137-1078/MA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3588
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  74 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-420-526-5

  Query Match             11.9%;  Score 52.5;  DB 1;  Length 74;
  Best Local Similarity   25.4%;  Pred. No. 15;
  Matches   17;  Conservative   10;  Mismatches   25;  Indels   15;  Gaps    2;

Qy      12 ILLLCFSIFSTEGK----------RRPAKAWSG-----RRTRLCCHRVPSPNSTNLKGHH 56
           :|||   :|:::|           |:|:  :||        :  |:|   ||   | |
Db       8 LLLLLLFVFASELNMVAEVQGATCRKPSMYFSGACFSDTNCQKACNREDWPNGKCLVGFK 67

Qy      57 VRLCKPC 63
             :||
Db      68 CECQRPC 74


RESULT 4
US-08-974-549A-174
; Sequence 174, Application US/08974549A
; Patent No. 6166178
;  GENERAL INFORMATION:
;    APPLICANT:  Cech, Thomas R.
;    APPLICANT:  Lingner, Joachim
;    APPLICANT:  Nakamura, Toru
;    APPLICANT:  Chapman, Karen B.
;    APPLICANT:  Morin, Gregg B.
;    APPLICANT:  Harley, Calvin B.
;    APPLICANT:  Andrews, William H.
;    TITLE OF INVENTION:  Human Telomerase Catalytic Subunit
;    NUMBER OF SEQUENCES:  727
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
```

```
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/974,549A
;       FILING DATE:  19-NOV-1997
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/724,643
;       FILING DATE:  01-OCT-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/844,419
;       FILING DATE:  18-APR-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/846,017
;       FILING DATE:  25-APR-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/851,843
;       FILING DATE:  06-MAY-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/854,050
;       FILING DATE:  09-MAY-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/911,312
;       FILING DATE:  14-AUG-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/912,951
;       FILING DATE:  14-AUG-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/915,503
;       FILING DATE:  14-AUG-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/US97/17618
;       FILING DATE:  01-OCT-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/US97/17885
;       FILING DATE:  01-OCT-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Apple, Randolph Ted
;       REGISTRATION NUMBER:  36,429
;       REFERENCE/DOCKET NUMBER:  015389-002610US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  174:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  42 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-974-549A-174

  Query Match          11.5%;  Score 51;  DB 4;  Length 42;
  Best Local Similarity  36.8%;  Pred. No. 12;
  Matches   14;  Conservative    4;  Mismatches   18;  Indels    2;  Gaps    1;

Qy       4 LVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCC 41
           |    | |:|||   ||  :  ||  |    |  ::||
Db       1 LTTDYLFCLLLLISYIF--KRSRRNEKKTKERDFKICC 36


RESULT 5
US-09-149-476-523
; Sequence 523, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
```

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 11.5%; Score 51; DB 4; Length 47;
Best Local Similarity 46.2%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

```
Qy       3 LLVLSSLLCILLLCFSIFSTEGKRRP 28
           || | | ||:||   |   |  |::|
Db      22 LLALGSXLCLLLCLPSFNKTRRKQKP 47
```

RESULT 6
US-08-637-759B-123
; Sequence 123, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA

```
;       ZIP:  30309-3450
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/637,759B
;       FILING DATE:  03-MAY-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/GB95/02875
;       FILING DATE:  11-DEC-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pabst, Patrea L.
;       REGISTRATION NUMBER:  31,284
;       REFERENCE/DOCKET NUMBER:  RPMS 101
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (404) 873-8794
;       TELEFAX:  (404) 873-8795
;   INFORMATION FOR SEQ ID NO:  123:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  56 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-08-637-759B-123

  Query Match             11.3%;  Score 50;  DB 2;  Length 56;
  Best Local Similarity   34.4%;  Pred. No. 22;
  Matches   11;  Conservative    4;  Mismatches   13;  Indels    4;  Gaps    1;

Qy      41 CHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLW 72
           | ||     :||  :| |    | |  : |:|
Db      14 CQRVTERLQSNLSRYHNR----CSLRSDVRVW 41


RESULT 7
US-08-871-355A-123
; Sequence 123, Application US/08871355A
; Patent No. 6015669
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/871,355A
;      FILING DATE:  09-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  56 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-871-355A-123

  Query Match             11.3%;  Score 50;  DB 3;  Length 56;
  Best Local Similarity   34.4%;  Pred. No. 22;
  Matches   11;  Conservative    4;  Mismatches   13;  Indels    4;  Gaps    1;

Qy      41 CHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLW 72
           | ||     :||  :| |    | |  : |:|
Db      14 CQRVTERLQSNLSRYHNR----CSLRSDVRVW 41


RESULT 8
US-09-201-945-123
; Sequence 123, Application US/09201945
; Patent No. 6342215
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/201,945
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/637,759
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  56 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-201-945-123

  Query Match             11.3%;  Score 50;  DB 4;  Length 56;
  Best Local Similarity   34.4%;  Pred. No. 22;
  Matches   11;  Conservative    4;  Mismatches   13;  Indels    4;  Gaps    1;

Qy      41 CHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLW 72
```

```
          | ||     :||  :| |    | |  : |:|
Db      14 CQRVTERLQSNLSRYHNR----CSLRSDVRVW 41


RESULT 9
US-08-726-306A-174
; Sequence 174, Application US/08726306A
; Patent No. 5958684
;  GENERAL INFORMATION:
;    APPLICANT:  van Leeuwen, Frederik Willem
;    APPLICANT:  Burbach, Johannes Peter Henri
;    APPLICANT:  Grosveld, Franklin G.
;    TITLE OF INVENTION:  DIAGNOSIS METHOD AND REAGENTS
;    NUMBER OF SEQUENCES:  189
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1 Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,306A
;      FILING DATE:  02-Oct-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 95/20080.4
;      FILING DATE:  02-Oct-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/009,832
;      FILING DATE:  01-Jan-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen M.
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  96,048-A (3255/00784)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 345-9100
;      TELEFAX:  (617) 345-9111
;  INFORMATION FOR SEQ ID NO:  174:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-726-306A-174

  Query Match             11.3%;  Score 50;  DB 2;  Length 77;
  Best Local Similarity   24.6%;  Pred. No. 32;
  Matches   16;  Conservative    6;  Mismatches   13;  Indels   30;  Gaps    3;

Qy      25 KRRP-AKAWSGRRTRLCCH--------------------RVPSPNSTNLKGHHVRLCKP 62
           :|||  :||: || |  |                    | |:|          | |:|
Db      21 RRRPRRRAWTCRRRRRRCRRSAATCGATTRKRWASCSARSRAPAPRR--------RRCRP 72

Qy      63 CKLEP 67
             :  |
Db      73 RRATP 77


RESULT 10
US-08-443-568B-10
; Sequence 10, Application US/08443568B
; Patent No. 5759807
;  GENERAL INFORMATION:
;    APPLICANT:  Breece, Tim
;    APPLICANT:  Hayenga, Kirk
;    APPLICANT:  Rindersknecht, Ernst
;    APPLICANT:  Vandlen, Richard
;    APPLICANT:  Daniel, Yansura
;    TITLE OF INVENTION:  Process for Producing Relaxin
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/443,568B
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/080,354
;      FILING DATE:  21-JUNE-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Abrams, Samuel B.
;      REGISTRATION NUMBER:  30,605
;      REFERENCE/DOCKET NUMBER:  7842-037
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-443-568B-10

  Query Match             11.2%;  Score 49.5;  DB 1;  Length 77;
  Best Local Similarity   37.9%;  Pred. No. 37;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    7;  Gaps    1;

Qy      21 STEGKRRPAKAWSGRRTRL-------CCH 42
           ||  ||:|    ||:: :|       |||
Db      37 STWSKRKPTGYGSGKKRQLYSALANKCCH 65


RESULT 11
PCT-US94-06997-10
; Sequence 10, Application PC/TUS9406997
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  460 Point San Bruno Boulevard
;    APPLICANT:  South San Francisco, California
;    APPLICANT:  United States of America
;    TITLE OF INVENTION:  Process for Producing Relaxin
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/06997
;      FILING DATE:  20-JUN-1994
```

```
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Halluin, Albert P.
;      REGISTRATION NUMBER:  25,227
;      REFERENCE/DOCKET NUMBER:  7842-025-228
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-854-3660
;      TELEFAX:  415-854-3694
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US94-06997-10

  Query Match             11.2%;  Score 49.5;  DB 5;  Length 77;
  Best Local Similarity   37.9%;  Pred. No. 37;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    7;  Gaps    1;

Qy      21 STEGKRRPAKAWSGRRTRL-------CCH 42
           ||  ||:|    ||:: :|       |||
Db      37 STWSKRKPTGYGSGKKRQLYSALANKCCH 65


RESULT 12
5256770-8
;Patent No. 5256770
;    APPLICANT: GLASER, CHARLES B.;MORSER, MICHAEL J.;LIGHT,
;DAVID R.
;    TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
;    NUMBER OF SEQUENCES: 48
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/506,325
;      FILING DATE: 09-APR-1990
;SEQ ID NO:8:
;      LENGTH: 36
5256770-8

  Query Match             10.9%;  Score 48;  DB 6;  Length 36;
  Best Local Similarity   34.5%;  Pred. No. 23;
  Matches   10;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       9 LLCILLLCFSIFSTEGKRRPAKAWSGRRT 37
           | |:|||| ::| :  :   |:   | |:
Db       8 LCCVLLLCGAVFVSPSEEIHARFRRGARS 36


RESULT 13
5244676-2
;Patent No. 5244676
;    APPLICANT: BELL, LESLIE D.;MAYER, ERNEST J.;PALMIER, MARK O.
;TOLUNAY, H.ESER;WARREN, THOMAS G.;WUN, TZE-CHEIN
;    TITLE OF INVENTION: MODIFIED TISSUE PLASMINOGEN ACTIVATOR
;WITH MODIFIED GLYCOSYLATION SITE
;    NUMBER OF SEQUENCES: 9
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/203,047
;      FILING DATE: 06-JUN-1988
;SEQ ID NO:2:
;      LENGTH: 40
5244676-2

  Query Match             10.9%;  Score 48;  DB 6;  Length 40;
  Best Local Similarity   34.5%;  Pred. No. 26;
  Matches   10;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       9 LLCILLLCFSIFSTEGKRRPAKAWSGRRT 37
           | |:|||| ::| :  :   |:   | |:
Db       8 LCCVLLLCGAVFVSPSQEIHARFRRGARS 36


RESULT 14
US-08-743-975-9
; Sequence 9, Application US/08743975
; Patent No. 6057434
;  GENERAL INFORMATION:
;    APPLICANT:  Jian Ni, Guo-Liang Yu and Reiner Gentz
;    TITLE OF INVENTION:  Mammary Transforming Protein
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/743,975
;      FILING DATE:  01 NOVEMBER 1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/006,187
;      FILING DATE:  02 NOVEMBER 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-507 (PF212)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-743-975-9

  Query Match             10.9%;  Score 48;  DB 3;  Length 61;
  Best Local Similarity   28.3%;  Pred. No. 43;
  Matches   13;  Conservative    7;  Mismatches   16;  Indels   10;  Gaps    2;

Qy      17 FSIFSTEGKRRPAKAWSGR------RTRLCCHRVPSPNSTNLKGHH 56
           | :||      || : | |       | :|    : |||  ::  :|
Db      13 FVVFSINLLSRPEREWEGMPQKGSGRAKL----LQSPNRKHISTNH 54


RESULT 15
US-07-689-693B-11
; Sequence 11, Application US/07689693B
; Patent No. 5231011
;  GENERAL INFORMATION:
;    APPLICANT:  David Hillyard
;    APPLICANT:  Baldomero M. Olivera
;    TITLE OF INVENTION:  Segregated Folding Determinants
;    TITLE OF INVENTION:  for Small Disulfide-Rich Peptides
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Thorpe, No. 5231011th & Western
;      STREET:  9035 South 700 East, Suite 200
;      CITY:  Sandy
;      STATE:  Utah
;      COUNTRY:  USA
;      ZIP:  84070
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 720 Kb storage
;      COMPUTER:  Compaq LTE/286
```

```
;      OPERATING SYSTEM:  DOS 4.01
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/689,693B
;      FILING DATE:  19910418
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  na
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Western, M. Wayne
;      REGISTRATION NUMBER:  22,788
;      REFERENCE/DOCKET NUMBER:  9925
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (801) 566-6633
;      TELEFAX:  (801) 566-0750
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Prepropeptide sequence for three loop
;      NAME/KEY:  conotoxin from Circler conotoxin from Conus
;      NAME/KEY:  textile
;      IDENTIFICATION METHOD:  Libraries were created
;      IDENTIFICATION METHOD:  using oligo-dT primed pUC13 vector
US-07-689-693B-11

  Query Match             10.7%;  Score 47.5;  DB 1;  Length 68;
  Best Local Similarity   27.4%;  Pred. No. 57;
  Matches   20;  Conservative   10;  Mismatches   18;  Indels   25;  Gaps    5;

Qy       6 LSSLLCILLLCFSIFST--EGKR---RPAKAWSGR----------RTRLCCHRVPSPNST 50
           | :|| | || ||: :   :| :   :||:    |           : ||     | :
Db       4 LGALLTICLLLFSLTAVPLDGDQHADQPAQRLQDRIPTEDHPLFDPNKRCC----PPVAC 59

Qy      51 NLKGHHVRLCKPC 63
           |:       ||||
Db      60 NMG------CKPC 66


Search completed: May  8, 2003, 21:48:40
Job time : 19 secs
```

GenCore version 5.1.5

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2003, 21:48:13 ; Search time 1349 Seconds
(without alignments)
972.450 Million cell updates/sec

Title: US-09-599-087B-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 227956

Minimum DB seq length: 25
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09599087/runat_07052003_113321_7356/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=25 -MAXLEN=80
-USER=US09599087_@CGN_1_1_1525_@runat_07052003_113321_7356 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	50.5	11.4	60	17	AZ918344	AZ918344 1006004B0
c	2	50.5	11.4	69	17	BH639950	BH639950 1008032H0
c	3	49	11.1	72	14	BQ394194	BQ394194 NISC_ng07
	4	49	11.1	76	17	AZ918366	AZ918366 1006004B0
c	5	49	11.1	79	9	AI877052	AI877052 uc56b11.r
	6	48	10.9	80	10	AV676540	AV676540 AV676540
c	7	47	10.6	67	17	AZ605110	AZ605110 1M0426I15
	8	47	10.6	73	10	AW600148	AW600148 SWL4CAK10
c	9	47	10.6	74	9	AA739046	AA739046 vv66d04.r
c	10	47	10.6	74	13	BM283913	BM283913 ki35b04.y
c	11	46.5	10.5	79	9	AI468972	AI468972 ti43h02.x
	12	46	10.4	50	9	AU105245	AU105245 AU105245
c	13	46	10.4	70	14	BQ092063	BQ092063 fz36f08.x
c	14	46	10.4	70	14	BQ092271	BQ092271 fz39e09.x
c	15	46	10.4	73	14	H22549	H22549 yn69h12.r1
	16	46	10.4	77	9	AA049559	AA049559 mj35b12.r
c	17	45.5	10.3	77	14	BQ666094	BQ666094 pb35g10.y
c	18	45.5	10.3	77	14	BQ811778	BQ811778 1030025E1
	19	45	10.2	67	13	BM518515	BM518515 ki93f08.y
c	20	45	10.2	70	17	AZ992067	AZ992067 2M0276O01
	21	45	10.2	72	14	BQ757145	BQ757145 EBem10_SQ
c	22	45	10.2	72	17	AZ493456	AZ493456 1M0328G07
c	23	45	10.2	72	17	AZ921678	AZ921678 1006031C0
	24	45	10.2	73	9	AA790708	AA790708 vw18f08.r
	25	45	10.2	79	9	AA498761	AA498761 vh32d12.r
	26	45	10.2	80	9	AU257386	AU257386 AU257386
	27	44	10.0	74	13	BM097862	BM097862 EBem04_SQ
	28	44	10.0	76	9	AU244358	AU244358 AU244358
	29	44	10.0	78	6	BQ241316	Bq241316 TaE05006F
	30	43.5	9.8	66	17	BH620911	BH620911 1007099G0
c	31	43.5	9.8	67	17	AZ808082	AZ808082 2M0071I24
c	32	43.5	9.8	72	14	R33083	R33083 yh77c06.s1
	33	43	9.7	40	9	AI589916	AI589916 tm81h06.x
	34	43	9.7	48	17	TA176E12P	AL475046 T. brucei
	35	43	9.7	51	10	AW232755	AW232755 fj30a01.x
c	36	43	9.7	58	9	AA948249	AA948249 oq33g08.s
	37	43	9.7	58	17	AZ835896	AZ835896 2M0130C22
c	38	43	9.7	62	17	FR0012583	AL003835 F.rubripe
	39	43	9.7	64	10	AW626440	AW626440 SWOvAFCAP
	40	43	9.7	64	10	AW626725	AW626725 SWOvAFCAP
	41	43	9.7	64	12	BF118532	BF118532 SWOvL3CAN
	42	43	9.7	70	9	AU258504	AU258504 AU258504
c	43	43	9.7	72	10	AV911900	AV911900 AV911900
c	44	43	9.7	75	9	AL643544	AL643544 AL643544
c	45	43	9.7	75	13	BI753134	BI753134 603025986

ALIGNMENTS

RESULT 1
AZ918344/c
LOCUS AZ918344 60 bp DNA linear GSS 17-DEC-2001
DEFINITION 1006004B02.x1 1006 - RescueMu Grid G Zea mays genomic, DNA sequence.
ACCESSION AZ918344
VERSION AZ918344.1 GI:13387628
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 60)
AUTHORS Walbot,V.

```
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006004  row: 36
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        5 a     17 c     26 g     12 t
ORIGIN

Alignment Scores:
Pred. No.:              9.11e+03        Length:         60
Score:                  50.50           Matches:        10
Percent Similarity:     78.57%          Conservative:   1
Best Local Similarity:  71.43%          Mismatches:     2
Query Match:            11.43%          Indels:         1
DB:                     17              Gaps:           1

US-09-599-087B-5 (1-81) x AZ918344 (1-60)

Qy      34 GlyArgArgThrArgLeuCysCysHisArgValProSerPro 47
           |||   ||||||||||||   |||||||||   |||:::|||
Db      43 GGCCACCGCACACGGCTCATCTGTCACCGG---CCCACGCCA 5

RESULT 2
BH639950/c
LOCUS       BH639950                 69 bp    DNA     linear   GSS 14-FEB-2002
DEFINITION  1008032H07.2EL_y1 1008 - RescueMu Grid I Zea mays genomic, DNA
            sequence.
ACCESSION   BH639950
VERSION     BH639950.1  GI:18665756
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008032  row: 22
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1008 - RescueMu Grid I"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        7 a     30 c     26 g      6 t
ORIGIN

Alignment Scores:
Pred. No.:              1.11e+04        Length:         69
Score:                  50.50           Matches:        10
Percent Similarity:     66.67%          Conservative:   2
Best Local Similarity:  55.56%          Mismatches:     5
Query Match:            11.43%          Indels:         1
DB:                     17              Gaps:           1

US-09-599-087B-5 (1-81) x BH639950 (1-69)

Qy      25 LysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHis 42
           :::||||||||||||:::   |||      |||   ||||||   |||   |||
Db      62 CGCCGGAGGCCGGCGCGGCGGTGGCGC---CGGCTGACCCGGAGTTGCCTCCAC 12

RESULT 3
BQ394194/c
LOCUS       BQ394194                 72 bp    mRNA    linear   EST 22-MAY-2002
DEFINITION  NISC_ng07g11.x1 NICHD XGC Emb6 Silurana tropicalis cDNA clone
            IMAGE:5382740 3', mRNA sequence.
ACCESSION   BQ394194
VERSION     BQ394194.1  GI:21081881
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   NIH-XCG http://image.llnl.gov/image/html/xenopuslib_info.shtml.
  TITLE     National Institute of Child Health and Human Development, National
            Cancer Insitute, Xenopus Gene Collection
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11974  row: N  column: 21
            Seq primer: -21M13 forward primer (ABI).
```

```
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:5382740"
                     /clone_lib="NICHD XGC Emb6"
                     /tissue_type="neurula"
                     /dev_stage="embryo, stages 14-19"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6.ccdb; Site_1: NotI; Site_2:
                     EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average
                     insert size 2.1 kb. Constructed by Invitrogen. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT       13 a     14 c      5 g     40 t
ORIGIN

Alignment Scores:
Pred. No.:              1.7e+04        Length:        72
Score:                  49.00          Matches:       9
Percent Similarity:     65.22%         Conservative:  6
Best Local Similarity:  39.13%         Mismatches:    8
Query Match:            11.09%         Indels:        0
DB:                     14             Gaps:          0

US-09-599-087B-5 (1-81) x BQ394194 (1-72)

Qy      17 PheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArg 36
           ||||||:::|||      :::||||||:::::      |||         |||   |||
Db      69 TTTAGTGTTTTTGAAAAAAAAGGGAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGC 10

Qy      37 ThrArgLeu 39
           :::|||::: 
Db       9 TCTAGAGTA 1

RESULT 4
AZ918366
LOCUS       AZ918366                 76 bp    DNA     linear   GSS 17-DEC-2001
DEFINITION  1006004B05.2EL_x3 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ918366
VERSION     AZ918366.1  GI:13387650
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006004  row: 36
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       15 a     24 c      22 g     15 t
ORIGIN

Alignment Scores:
Pred. No.:              1.84e+04       Length:        76
Score:                  49.00          Matches:       10
Percent Similarity:     57.89%         Conservative:  1
Best Local Similarity:  52.63%         Mismatches:    2
Query Match:            11.09%         Indels:        6
DB:                     17             Gaps:          1

US-09-599-087B-5 (1-81) x AZ918366 (1-76)

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           |||   :::||||||||||                    ||||||||||||||||||
Db      38 CACTCAATGCCATCGCCA----------------CCCCACCACGTCAGATTATGT 76

RESULT 5
AI877052/c
LOCUS       AI877052                 79 bp    mRNA    linear   EST 21-JUL-1999
DEFINITION  uc56b11.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:1429629 5' similar to TR:O35500 O35500 KINESIN LIGHT CHAIN
            ISOFORM D. ;, mRNA sequence.
ACCESSION   AI877052
VERSION     AI877052.1  GI:5551101
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP);
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:913697
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1429629"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
```

rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 15 a 23 c 25 g 16 t
ORIGIN

Alignment Scores:

Pred. No.:	1.95e+04	Length:	79
Score:	49.00	Matches:	10
Percent Similarity:	47.83%	Conservative:	1
Best Local Similarity:	43.48%	Mismatches:	12
Query Match:	11.09%	Indels:	0
DB:	9	Gaps:	0

US-09-599-087B-5 (1-81) x AI877052 (1-79)

```
Qy      40 CysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeu 59
           |||         |||||||||   :::|||   |||      |||        |||
Db      78 TGCGCAGGAGCAGTTCCTTCTCTAAGCTCCGGGAATCCATCAGGCGAGGAAGCGAGAAGC 19

Qy      60 CysLysPro 62
           |||   |||
Db      18 TGTTCTCCC 10
```

RESULT 6
AV676540
LOCUS AV676540 80 bp mRNA linear EST 05-OCT-2000
DEFINITION AV676540 Nori Satoh unpublished cDNA library Ciona intestinalis cDNA clone rcitble23 3', mRNA sequence.
ACCESSION AV676540
VERSION AV676540.1 GI:10114539
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 80)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES Location/Qualifiers
source 1. .80
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="rcitble23"
/clone_lib="Nori Satoh unpublished cDNA library"
/tissue_type="whole animal"
/dev_stage="tailbud"
/note="Vector: pBluescript SK"
BASE COUNT 20 a 19 c 15 g 26 t
ORIGIN

Alignment Scores:

Pred. No.:	2.53e+04	Length:	80
Score:	48.00	Matches:	8
Percent Similarity:	68.42%	Conservative:	5
Best Local Similarity:	42.11%	Mismatches:	4
Query Match:	10.86%	Indels:	2
DB:	10	Gaps:	1

US-09-599-087B-5 (1-81) x AV676540 (1-80)

```
Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArg------ValProSerProAsn 48
           |||   |||:::|||      ||||||   ::::::      :::||||||||:::
Db      15 TGGGGAGGTAAAAGATGCTCCTTGTGTTCGTACCAATATGTTATTCCAAGCCCAGAT 71
```

RESULT 7
AZ605110/c
LOCUS AZ605110 67 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0426I15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0426I15 F, DNA sequence.
ACCESSION AZ605110
VERSION AZ605110.1 GI:11727300
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 67)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0426 row: I column: 15
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 67.
FEATURES Location/Qualifiers
source 1. .67
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0426I15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 15 a 17 c 22 g 13 t
ORIGIN

Alignment Scores:

Pred. No.:	2.49e+04	Length:	67
Score:	47.00	Matches:	8
Percent Similarity:	70.59%	Conservative:	4
Best Local Similarity:	47.06%	Mismatches:	5
Query Match:	10.63%	Indels:	0
DB:	17	Gaps:	0

US-09-599-087B-5 (1-81) x AZ605110 (1-67)

```
Qy      44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||  ||||||  :::::::::  |||  |||  :::|||||||
Db      66 GTGCCACCCCCCAACCCCTCTTCTGTGCTAGGGGTCCATTCAAAGCTTTGT 16

RESULT 8
AW600148
LOCUS       AW600148                 73 bp    mRNA    linear   EST 22-MAR-2000
DEFINITION  SWL4CAK10E07SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SWL4CAK10E07 5', mRNA sequence.
ACCESSION   AW600148
VERSION     AW600148.1  GI:7287661
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SWL4CAK10E07"
                     /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                     /dev_stage="larval stage four"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from L4s isolated from the peritoneal
                     cavity of jirds and converted to double-stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNA pol I. The library has 2.7 x 105
                     independent recombinants and the average insert size is
                     approx. 1050bp. The library was constructed by Michelle
                     Lizotte-Waniewski. The library is available from Dr. S.A.
                     Williams, email:genome@neal.smith.edu."
BASE COUNT       21 a     11 c     10 g     31 t
ORIGIN

Alignment Scores:
Pred. No.:              2.82e+04        Length:         73
Score:                  47.00           Matches:        10
Percent Similarity:     70.59%          Conservative:   2
Best Local Similarity:  58.82%          Mismatches:     5
Query Match:            10.63%          Indels:         0
DB:                     10              Gaps:           0

US-09-599-087B-5 (1-81) x AW600148 (1-73)

Qy       6 LeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThr 22
           ||||||   :::   ||||||   ||||||   |||||||:::|||   |||
Db       3 CTATCACTAGTGTTTTGCATTAGTTTGCTATCTTTTTCAATGTTTGAAACA 53

RESULT 9
AA739046/c
LOCUS       AA739046                 74 bp    mRNA    linear   EST 14-JAN-1998
DEFINITION  vv66d04.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1227367 5' similar to gb:M73436 Mouse ribosomal protein S4
            (MOUSE);, mRNA sequence.
ACCESSION   AA739046
VERSION     AA739046.1  GI:2775852
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:652959
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1227367"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     24 c     15 g     16 t
ORIGIN

Alignment Scores:
Pred. No.:              2.88e+04        Length:         74
Score:                  47.00           Matches:        8
Percent Similarity:     83.33%          Conservative:   2
Best Local Similarity:  66.67%          Mismatches:     2
Query Match:            10.63%          Indels:         0
DB:                     9               Gaps:           0

US-09-599-087B-5 (1-81) x AA739046 (1-74)

Qy      47 ProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||||||||:::::   ||||||   ||||||
Db      54 CCGAATAGTACGAGCATCATGGGTCACCAGGTGCGG 19

RESULT 10
BM283913/c
LOCUS       BM283913                 74 bp    mRNA    linear   EST 28-DEC-2001
DEFINITION  ki35b04.y1 Ascaris suum female gonad GZ SL1 TOPO v1 Murphy
            Chiapelli McCarter Ascaris suum cDNA 5', mRNA sequence.
ACCESSION   BM283913
VERSION     BM283913.1  GI:17992955
KEYWORDS    EST.
SOURCE      pig roundworm.
  ORGANISM  Ascaris suum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
            ; Ascarididae; Ascaris.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
```

```
          , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
          ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
          , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
          Wilson,R.
  TITLE   The Washington Univ. Nematode EST Project, 1999
  JOURNAL Unpublished (1999)
COMMENT   Contact: McCarter JP
          The Washington Univ. Nematode EST Project, 1999
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          The library was constructed by Claire Murphy, Brandi Chiapelli, and
          Dr. James McCarter at Washington University, St. Louis. DNA
          Sequencing by: Washington University Genome Sequencing Center
          Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Ascaris suum"
                     /db_xref="taxon:6253"
                     /clone_lib="Ascaris suum female gonad GZ SL1 TOPO v1
                     Murphy Chiapelli McCarter"
                     /sex="Female"
                     /tissue_type="Dissected female gonad (GZ=germinal zone,
                     most distal region of gonad)"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy, Brandi Chiapelli, and Dr. James McCarter at
                     Washington University, St. Louis. Oligo(dT)-SL1 PCR based
                     library. Ascaris suum female gonad GZ cDNA PCR products of
                     size >400 nucleotides containing SL1 on the 5' end and
                     oligo(dT) on the 3' end were non-directionally cloned
                     into pCRII-TOPO(Invitrogen) following the Topo TA cloning
                     protocol. Dissected nematode tissues were provided by Dr.
                     Alan Scott (ascott@jhsph.edu) of the School of Public
                     Hygiene and Public Health at John Hopkins University in
                     Baltimore, MD."
BASE COUNT      30 a     9 c     22 g     13 t
ORIGIN

Alignment Scores:
Pred. No.:              2.88e+04        Length:         74
Score:                  47.00           Matches:        10
Percent Similarity:     73.33%          Conservative:   1
Best Local Similarity:  66.67%          Mismatches:     4
Query Match:            10.63%          Indels:         0
DB:                     13              Gaps:           0

US-09-599-087B-5 (1-81) x BM283913 (1-74)

Qy       9 LeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThrGlu 23
           |||   |||   ||||||||||||||   |||   |||:::|||
Db      61 CTCCACTGCTTTCTCCTCCTTTGCTTCTTCATCACATCTTCGGAA 17

RESULT 11
AI468972/c
LOCUS       AI468972                79 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  ti43h02.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2133267 3'
            similar to WP:F59E12.9 CE11534 ;, mRNA sequence.
ACCESSION   AI468972
VERSION     AI468972.1  GI:4331062
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11547-015
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2133267"
                     /clone_lib="NCI_CGAP_Lym12"
                     /tissue_type="lymphoma, follicular mixed small and large
                     cell"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.25 kb. Life Technologies
                     catalog #: 11547-015"
BASE COUNT      12 a     35 c     31 g      1 t
ORIGIN

Alignment Scores:
Pred. No.:              3.57e+04        Length:         79
Score:                  46.50           Matches:        10
Percent Similarity:     54.55%          Conservative:   2
Best Local Similarity:  45.45%          Mismatches:     7
Query Match:            10.52%          Indels:         3
DB:                     9               Gaps:           1

US-09-599-087B-5 (1-81) x AI468972 (1-79)

Qy      58 ArgLeuCysLysProCysLys---------LeuGluProGluProArgLeuTrpValVal 74
           |||   |||   |||                :::   |||   ||||||:::||||||
Db      77 CGTTGGTGTGGGCCCCCCCCCAGGGGGGGGGTGGCCCCCCCCCCGCGTGTTTGGGTGCCC 18

Qy      75 ProGly 76
           ||||||
Db      17 CCCGGG 12

RESULT 12
AU105245
LOCUS       AU105245                50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105245 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            COL05671, mRNA sequence.
ACCESSION   AU105245
VERSION     AU105245.1  GI:13554766
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
```

a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL05671"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"
BASE COUNT 2 a 19 c 18 g 11 t
ORIGIN

Alignment Scores:

Pred. No.:	2.08e+04	Length:	50
Score:	46.00	Matches:	9
Percent Similarity:	73.33%	Conservative:	2
Best Local Similarity:	60.00%	Mismatches:	4
Query Match:	10.41%	Indels:	0
DB:	9	Gaps:	0

US-09-599-087B-5 (1-81) x AU105245 (1-50)

```
Qy      32 TrpSerGlyArgArgThrArgLeuCysCysHisArgValProSer 46
           :::|||   ||||||   ||||||   |||   |||:::||||||
Db       5 TTTTCCGCCCGGCGCTGTCGGCTGCTGTGCCCGCGGCTCCCCAGT 49
```

RESULT 13
BQ092063/c
LOCUS BQ092063 70 bp mRNA linear EST 08-APR-2002
DEFINITION fz36f08.x1 zebrafish fin day1 regeneration Danio rerio cDNA clone 5905287 3', mRNA sequence.
ACCESSION BQ092063
VERSION BQ092063.1 GI:20072973
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 70)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)
Seq primer: T7 from Gibco.
FEATURES Location/Qualifiers
source 1. .70
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5905287"
/clone_lib="zebrafish fin day1 regeneration"
/sex="mixed male and female"
/tissue_type="1 day fin regenerates"
/lab_host="E. coli XLOLR"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; 1st strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed by second strand synthesis, and ligated to 5' adapter (5')-aattcggcacgag- 3', 3'-gccgtgctc-5'. cDNA was cloned directionally (EcoRI/XhoI) into Stratagene Zap express lambda phage arms. Mass invivo excision done to obtain inserts in pBK-CMV phagemid."
BASE COUNT 5 a 19 c 11 g 35 t
ORIGIN

Alignment Scores:

Pred. No.:	3.38e+04	Length:	70
Score:	46.00	Matches:	8
Percent Similarity:	66.67%	Conservative:	2
Best Local Similarity:	53.33%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	14	Gaps:	0

US-09-599-087B-5 (1-81) x BQ092063 (1-70)

```
Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
           |||:::   ||||||         ||||||   :::||||||||||
Db      63 GGAAGAGAGAGGCCACAAACACAGTGGTCGTTCAAGAGGACAAGA 19
```

RESULT 14
BQ092271/c
LOCUS BQ092271 70 bp mRNA linear EST 08-APR-2002
DEFINITION fz39e09.x1 zebrafish fin day1 regeneration Danio rerio cDNA clone 5905624 3', mRNA sequence.
ACCESSION BQ092271
VERSION BQ092271.1 GI:20073488
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 70)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)
Seq primer: T7 from Gibco.
FEATURES Location/Qualifiers
source 1. .70
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5905624"
/clone_lib="zebrafish fin day1 regeneration"
/sex="mixed male and female"
/tissue_type="1-day fin regenerates"
/lab_host="E. coli XLOLR"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; 1st strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed by second strand synthesis, and ligated to 5' adapter (5')-aattcggcacgag- 3', 3'-gccgtgctc-5'. cDNA was cloned

directionally (EcoRI/XhoI) into Stratagene Zap express lambda phage arms. Mass invivo excision done to obtain inserts in pBK-CMV phagemid."
BASE COUNT 5 a 19 c 11 g 35 t
ORIGIN

Alignment Scores:

Pred. No.:	3.38e+04	Length:	70
Score:	46.00	Matches:	8
Percent Similarity:	66.67%	Conservative:	2
Best Local Similarity:	53.33%	Mismatches:	5
Query Match:	10.41%	Indels:	0
DB:	14	Gaps:	0

US-09-599-087B-5 (1-81) x BQ092271 (1-70)

```
Qy     24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
          |||:::   ||||||         ||||||   :::||||||||||
Db     63 GGAAGAGAGAGGCCACAAACACAGTGGTCGTTCAAGAGGACAAGA 19
```

RESULT 15
H22549/c
LOCUS H22549 73 bp mRNA linear EST 06-JUL-1995
DEFINITION yn69h12.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone IMAGE:173735 5' similar to gb:X56009 GUANINE NUCLEOTIDE-BINDING PROTEIN G(S), ALPHA SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION H22549
VERSION H22549.1 GI:891244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 73)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 755
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 755 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .73
/organism="Homo sapiens"
/db_xref="GDB:3835770"
/db_xref="taxon:9606"
/clone="IMAGE:173735"
/clone_lib="Soares adult brain N2b5HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: brain; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT 3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."
BASE COUNT 20 a 21 c 17 g 10 t 5 others
ORIGIN

Alignment Scores:

Pred. No.:	3.59e+04	Length:	73
Score:	46.00	Matches:	9
Percent Similarity:	47.37%	Conservative:	0
Best Local Similarity:	47.37%	Mismatches:	10
Query Match:	10.41%	Indels:	0
DB:	14	Gaps:	0

US-09-599-087B-5 (1-81) x H22549 (1-73)

```
Qy     16 CysPheSerIlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGly 34
          |||  |||      |||      |||        ||||||     |||||||||
Db     59 TGTTGNTCCAGATGCTCTTNAAGAGGTTCAGAGCCTCCTGCAGACGGTTGGTCTGGT 3
```

Search completed: May 8, 2003, 22:36:25
Job time : 1353 secs